

WILLIAM ROBSON SCHWARTZ

**SEGMENTAÇÃO DE IMAGENS BASEADA EM DEPENDÊNCIA
ESPACIAL UTILIZANDO CAMPO ALEATÓRIO DE MARKOV
ASSOCIADO COM CARACTERÍSTICAS DE TEXTURAS**

Dissertação apresentada como requisito parcial à
obtenção do grau de Mestre. Programa de Pós-
Graduação em Informática, Setor de Ciências
Exatas, Universidade Federal do Paraná.
Orientador: Prof. Dr. Hélio Pedrini

CURITIBA

2005

SUMÁRIO

LISTA DE FIGURAS	ii
LISTA DE ALGORITMOS	vii
GLOSSÁRIO	viii
RESUMO	ix
ABSTRACT	x
1 INTRODUÇÃO	1
1.1 Objetivos e Contribuições	3
1.2 Organização do Trabalho	4
2 TRABALHOS RELACIONADOS	6
2.1 Modelos Estocásticos de Imagens	7
2.2 Características de Texturas	15
2.2.1 Abordagem Estatística	17
2.2.2 Abordagem Baseada em Processamento de Sinais	30
2.2.3 Abordagem Geométrica	35
2.2.4 Abordagem Baseada em Modelos Paramétricos	43
2.3 Formulação Bayesiana	45
2.4 Segmentação Baseada em Dependência Espacial	50
3 METODOLOGIA PROPOSTA	60
3.1 Aspectos Gerais	60
3.2 Segmentação Inicial	62
3.3 Segmentação Baseada em Pixels	72
4 RESULTADOS EXPERIMENTAIS	80
4.1 Seleção das Características de Texturas	80
4.2 Avaliação da Metodologia Proposta	91
5 CONCLUSÕES	102
REFERÊNCIAS BIBLIOGRÁFICAS	104

LISTA DE FIGURAS

2.1	Disposição espacial das variáveis aleatórias no modelo <i>Markov Mesh</i>	9
2.2	Configurações para dependência local no modelo <i>Markov Mesh</i> , o círculo preenchido indica a variável aleatória $X_{a,b}$	10
2.3	Vizinhança em relação ao pixel central e cliques. (a) vizinhança de primeira ordem e segunda ordem, respectivamente; (b) cliques das respectivas vizinhanças.	12
2.4	Principais aplicações da análise de texturas. (a) imagem monocromática original; (b) mapa de classes resultante da classificação da imagem, indicando as texturas presentes; (c) bordas identificadas pelo processo de segmentação; (d) textura sintética gerada pela amostragem de um campo aleatório de Markov.	16
2.5	Medidas calculadas a partir do histograma. (a) coeficiente de assimetria; (b) curtose.	18
2.6	Composição da matriz de co-ocorrência. Cada elemento é composto pelo número de transições específicas entre os níveis de cinza.	20
2.7	Ângulos utilizados para o cálculo das matrizes de co-ocorrência. As transições para cada ângulo são computadas sempre em relação ao pixel localizado na posição central.	21
2.8	Relações entre os canais de cores, onde os números iguais indicam as variações de intensidades consideradas, sempre passando pelo pixel central do canal H.	23
2.9	Ângulos utilizados para construção das matrizes de comprimento de corridas de cinza.	26
2.10	Ilustração da função de autocorrelação para uma textura composta por $M \times N$ pixels.	30
2.11	Função de autocorrelação para $\rho_{ff}(p, 0)$	30
2.12	Funções $S(\theta)$ determinadas a partir do espectro de Fourier das texturas mostradas em (a) e (c). O eixo das abscissas representa o parâmetro θ	33
2.13	Decomposição resultante da aplicação da transformada wavelet. (a) decomposição em um nível; (b) decomposição em dois níveis.	34
2.14	Representação do espectro de textura. (a), (c) imagens originais; (b), (d) respectivos espectros de textura representados em escala logarítmica.	36
2.15	Vizinhança utilizada para determinação da unidade de textura.	37

2.16	Decomposição da unidade de textura. (a) estrutura original da unidade de textura; (a) matriz diagonal; (b) matriz cruz.	39
2.17	Conectividade sobre a unidade de textura. (a) conectividade-4 de primeira ordem; (b) conectividade-4 de segunda ordem.	40
2.18	Valores utilizados para determinação de α e β no método TFCM para análise de texturas. As linhas indicam o tipo da primeira varredura, as colunas indicam o tipo da segunda varredura.	41
2.19	Representação da vizinhança do pixel central com valor x_s	44
2.20	Estimação dos parâmetros α , $b_{1,1}$ e $b_{1,2}$ do MRF que modela as texturas. (a) $\alpha = -6.0$, $b_{1,1} = 3.0$ e $b_{1,2} = 3.0$; (b) $\alpha = -2.0$, $b_{1,1} = 0.08$ e $b_{1,2} = 3.0$. . .	45
2.21	Segmentação de imagens. (a) imagem original; (b) segmentação obtida após a primeira etapa; (c) segmentação baseada em pixels.	51
2.22	Processo de segmentação de imagens baseado em MRF e na formulação Bayesiana.	52
3.1	Diagrama ilustrando a segmentação inicial. Após a validação do agrupamento tem-se como resultado o número de classes e os parâmetros que as descrevem.	63
3.2	Aplicação da PCA, acarretando sobreposição entre as distribuições das medidas se escolhido o componente associado ao maior autovalor, neste caso w_1	65
3.3	Espaço de características bidimensional. (a) espaço após o agrupamento de dados, ainda considerando as regiões heterogêneas; (b) espaço após o refinamento proporcionado pelo histograma.	66
3.4	Amostragem das classes homogêneas. (a) imagem original; (b) regiões obtidas a partir do histograma; (c) regiões amostradas, utilizando 100 amostras para cada classe.	69
3.5	Validação do agrupamento. (a) imagem original; (b) valores resultantes para cada um dos critérios.	72
3.6	Diagrama ilustrando a segmentação baseada em pixels. A segmentação final é obtida a partir da realização do MRF que modela a distribuição dos rótulos das classes.	73

3.7	Realização inicial do MRF para a imagem mostrada na figura 3.4. (a) inicialização aleatória; (b) inicialização com o classificador quadrático, desconsiderando a probabilidade <i>a priori</i> ; (b) regiões contendo classes homogêneas com inicialização aleatória; (d) regiões contendo classes homogêneas com o classificador quadrático, desconsiderando a probabilidade <i>a priori</i>	77
3.8	Segmentação final. (a) imagem original; (b) MRF obtido após o ICM; (c) fronteiras de classes indicadas sobre a imagem original.	79
4.1	Imagens utilizadas para avaliação das características de texturas. Conjunto <i>Brodatz</i> , composto pelas imagens D100, D105, D107, D10, D15, D26, D31, D35, D40, D59, D63, D72, D79, D88 e D91.	81
4.2	Imagens utilizadas para avaliação das características de texturas. Conjunto <i>MeasTex</i> , composto pelas imagens Concrete.0000, Concrete.0002, Concrete.0006, Concrete.0007, Concrete.0010, Grass.0000, Grass.0002, Grass.0008, Grass.0011, Grass.0016, Rock.0004, Rock.0007, Rock.0011, Rock.0019 e Rock.0023.	82
4.3	Classificação utilizando os próprios tons de cinza como característica de texturas, para variáveis padronizadas (Vap), variáveis sem padronização (Van), cinco primeiros componentes principais (P05) e para os dez primeiros componentes principais (P10). (a) classificação correta com variação no tamanho da janela; (b) tempo de execução conforme o tamanho da janela.	84
4.4	Classificação efetuada com alterações na ordem da vizinhança do método baseado no modelo generalizado de Ising. (a) variáveis padronizadas; (b) variáveis sem padronização.	85
4.5	Classificação efetuada com alterações no parâmetro Δ do método de codificação de características de texturas. (a) variáveis padronizadas; (b) variáveis sem padronização.	85
4.6	Classificação efetuada com alterações nos parâmetros p e q da função de autocorrelação. (a) variáveis padronizadas; (b) variáveis sem padronização. . .	85
4.7	Classificação efetuada com alterações no parâmetros do método de decomposição de valor singular, utilizando-se variáveis padronizadas. (a) tamanho da janela utilizada para decomposição; (b) número de valores singulares utilizados. .	86

4.8	Classificação baseada em matrizes de co-ocorrência, utilizando-se variáveis padronizadas. (a) alteração no parâmetro d ; (b) variação no número de tons de cinza presentes na imagem de entrada.	86
4.9	Classificação efetuada a partir de combinações distintas de tamanhos de janelas para classificação e treinamento, utilizando características extraídas a partir de matrizes de co-ocorrência.	87
4.10	Classificação efetuada a partir de combinações distintas de tamanhos de janelas para classificação e treinamento, utilizando características obtidas a partir da transformada wavelet.	87
4.11	Tempo requerido para classificação, utilizando os métodos de autocorrelação (Au), transformada wavelet (Da), estatísticas de primeira ordem (Fi), espectro de Fourier (Fo), matrizes de comprimento de corridas de cinza (Ga), unidade de textura (He), codificação de características de texturas (Ho), matrizes de co-ocorrência (Ha), decomposição de valor singular (Sv) e modelo generalizado de Ising (Is). (a) métodos rápidos; (b) métodos lentos.	88
4.12	Classificação obtida (utiliza-se a mesma legenda da figura 4.11). (a) variáveis padronizadas; (b) variáveis sem padronização.	88
4.13	Segmentação resultante para imagem do rio Nilo, composta de 512×512 pixels. (a) resultado considerando a dependência espacial; (b) resultado utilizando apenas o classificador quadrático; (c) realização resultante do MRF, considerando a dependência espacial; (d) regiões resultantes da segmentação baseada no classificador quadrático.	92
4.14	Segmentação resultante para o mosaico #1, composto de 512×512 pixels, contendo 5 classes distintas. (a) ICM com aplicação da perturbação, apresentando segmentação correta de 98.4%; (b) ICM sem aplicação da perturbação, a segmentação correta fica em 97.9%; (c)-(d) realizações resultantes do MRF.	93
4.15	Critérios para determinação do número de classes presentes na imagem da figura 4.14.	93
4.16	Segmentação resultante para o mosaico #2, composto de 512×512 pixels, contendo 5 classes distintas. (a) ICM com aplicação da perturbação, apresentando segmentação correta de 98.3%; (b) ICM sem aplicação da perturbação, a segmentação correta fica em 95.9%; (c)-(d) realizações resultantes do MRF.	94
4.17	Critérios para determinação do número de classes presentes na imagem da figura 4.16.	94

4.18	Segmentação resultante para o mosaico #3, composto de 512×512 pixels, contendo 4 classes distintas, embora dispostas em 5 regiões. (a) ICM com aplicação da perturbação, apresentando segmentação correta de 97.5%; (b) ICM sem aplicação da perturbação, a segmentação correta fica em 96.1%; (c)-(d) realizações resultantes do MRF.	95
4.19	Critérios para determinação do número de classes presentes na imagem da figura 4.18.	95
4.20	Segmentação resultante para o mosaico #4, composto de 256×256 pixels, contendo 5 classes distintas. (a) ICM com aplicação da perturbação, apresentando segmentação correta de 98.2%; (b) ICM sem aplicação da perturbação, a segmentação correta fica em 93.9%; (c)-(d) realizações resultantes do MRF.	96
4.21	Critérios para determinação do número de classes presentes na imagem da figura 4.20. (a) utilizando o método baseado em transformada wavelet; (b) utilizando método baseado em matrizes de comprimento de corridas de cinza.	96
4.22	Segmentação resultante para imagem composta de 512×512 pixels. (a) resultado considerando a dependência espacial; (b) resultado utilizando apenas o classificador quadrático; (c) realização resultante do MRF, considerando a dependência espacial; (d) regiões resultantes da segmentação baseada no classificador quadrático.	97
4.23	Critérios para determinação do número de classes presentes na imagem da figura 4.22.	97
4.24	Segmentação resultante para imagem composta de 512×256 pixels.	98
4.25	Critérios para determinação do número de classes presentes na imagem da figura 4.24.	98
4.26	Segmentação resultante para imagem composta de 640×640 pixels.	98
4.27	Alteração no número de execuções do <i>k-means</i> , cada curva representa um dos mosaicos apresentados anteriormente. (a) utilizando variáveis padronizadas; (b) utilizando variáveis sem padronização.	100
4.28	Decaimento da energia para o ICM, com perturbação (C_p) e sem perturbação (S_p). Considera-se, neste experimento, a energia total dividida pelo número de pixels contidos na imagem.	101

LISTA DE ALGORITMOS

2.1	Estimação MAP utilizando <i>simulated annealing</i>	47
2.2	Estimação MAP utilizando o método ICM.	48
2.3	Estimação MAP utilizando o método MPM.	49
3.1	Agrupamento de dados.	67
3.2	Histograma tridimensional.	69
3.3	Perturbação do ICM.	78

GLOSSÁRIO

BWS	<i>black-white symmetry</i>	simetria preto-branco
CDTM	<i>cross-diagonal texture matrix</i>	matriz de textura diagonal-e-cruz
DD	<i>degree of direction</i>	grau de direção
GLN	<i>gray level non-uniformity</i>	não uniformidade de tom de cinza
GLRLM	<i>gray level run length matrices</i>	matrizes de comprimento de corridas de cinza
GMRF	<i>Gaussian Markov random field</i>	campo aleatório de Markov Gaussiano
GRF	<i>Gibbs random field</i>	campo aleatório de Gibbs
GS	<i>geometric symmetry</i>	simetria geométrica
ICM	<i>iterated conditional modes</i>	modos condicionais de iteração
LBP	<i>local binary patterns</i>	padrões locais binários
LRE	<i>long runs emphasis</i>	ênfase em corridas longas
MDS	<i>micro-diagonal structure</i>	micro-estruturas orientadas na diagonal
MHS	<i>micro-horizontal structure</i>	micro-estruturas orientadas na horizontal
MPM	<i>maximizer of posterior marginals</i>	maximização de probabilidades marginais
MRF	<i>Markov random field</i>	campo aleatório de Markov
MVS	<i>micro-vertical structure</i>	micro-estruturas orientadas na vertical
NGRF	<i>neighbor Gibbs random field</i>	campo aleatório de Gibbs de vizinhança
PCA	<i>principal component analysis</i>	análise de componentes principais
RLN	<i>run length non-uniformity</i>	não uniformidade de tamanho de corrida
RP	<i>run percentage</i>	percentagem de corrida
SA	<i>simulated annealing</i>	resfriamento simulado
SM	<i>scatter matrices</i>	matrizes de espalhamento
SRE	<i>short runs emphasis</i>	ênfase em corridas curtas
SVD	<i>singular value decomposition</i>	decomposição de valor singular
TFCM	<i>texture feature coding method</i>	método para codificação de características de texturas
TS MRF	<i>tree-structured Markov random field</i>	campo aleatório de Markov estruturado por árvore
TU	<i>texture unit</i>	unidade de textura

RESUMO

Uma etapa crítica presente no processo de análise de imagens é a segmentação, responsável por obter informações de alto nível sobre as regiões ou objetos contidos na imagem, de modo a facilitar sua interpretação. Contudo, a segmentação ainda é um dos maiores desafios na área de análise de imagens, particularmente quando não se utiliza informações previamente adquiridas sobre a imagem a ser segmentada. Os métodos convencionais de segmentação desconsideram a dependência espacial entre as regiões, o que pode gerar resultados impróprios. Técnicas que consideram a dependência espacial entre as regiões da imagem têm recebido crescente atenção da comunidade científica, pois apresentam uma maior precisão nos resultados obtidos. Embora avanços significativos tenham sido alcançados na segmentação de texturas e de imagens coloridas separadamente, a combinação dessas duas propriedades é considerada como um problema bem mais complexo. Devido à importância dessa etapa no processo de análise de imagens e ao fato de não existirem soluções definitivas para o problema, este trabalho propõe o desenvolvimento de um novo método de segmentação aplicado a imagens texturizadas monocromáticas e coloridas. O método utiliza a formulação Bayesiana para associar a dependência espacial modelada por um campo aleatório de Markov com características de texturas. A segmentação final é obtida por meio da aplicação de técnicas de relaxação para minimizar uma função de energia definida a partir da referida associação. Experimentos são efetuados visando avaliar os métodos de análise de texturas, bem como validar a metodologia proposta.

ABSTRACT

A critical stage present in the image analysis process is the segmentation, responsible for obtaining high level information about regions or objects in the image, in order to facilitate its interpretation. However, the segmentation is still one of the greatest challenges in the image analysis area, particularly when it does not use information previously acquired on the image to be segmented. Conventional segmentation methods do not consider the spatial dependence between the regions, which can generate improper results. Techniques considering the spatial dependence between the image regions have received increasing attention from the scientific community, because they present a major precision in the obtained results. Although significant advances have been reached in the segmentation of textures and colored images separately, the combination of these two properties is considered a more complex problem. Due to the importance of this stage in the image analysis process and to the fact that does not exist definitive solutions to the problem, this work considers the development of a new segmentation method applied to gray scale and color texture images. The method uses the Bayesian formulation to associate the spatial dependence modeled by a Markov random field with texture features. The final segmentation is obtained by the application of relaxation techniques to minimize an energy function defined by such association. Experiments are performed to evaluate the texture analysis methods, as well as validating the proposal method.

CAPÍTULO 1

INTRODUÇÃO

Como a interpretação da grande quantidade de dados contidos em imagens digitais é uma atividade complexa, um processo intermediário de *segmentação* é necessário para particionar o conjunto de dados de entrada formando regiões homogêneas de modo a produzir estruturas de mais alto nível, correspondentes a objetos ou partes de objetos, que possam ser relacionados para viabilizar o processo de interpretação. Por isso, um processo de segmentação que identifique corretamente as formas, topologia e localização dos objetos é um requisito de fundamental importância para que as informações resultantes de um sistema de análise de imagens sejam confiáveis.

Contudo, a segmentação de imagens ainda é um dos maiores desafios na área de análise de imagens. Processar uma imagem de modo a segmentar um número de objetos, possivelmente em posições distintas e com diferentes tamanhos e formas, é uma tarefa difícil e extremamente dependente da correta extração de características dos objetos a partir de imagens ruidosas. Ruídos podem levar métodos de segmentação a distorcer as formas dos objetos, prejudicando seu reconhecimento: regiões distintas podem ser incorretamente identificadas como uma única região, ou uma região homogênea pode ser dividida em regiões menores.

Abordagens distintas têm sido propostas para efetuar a segmentação de imagens (Pal e Pal 1993, Reed e Dubuf 1993), sendo normalmente classificadas como *baseada em regiões*, *baseada em bordas* ou *híbridas* dessas duas. Métodos de segmentação por regiões agrupam pontos da imagem que apresentam valores semelhantes para uma determinada característica, cuja escolha é relacionada ao domínio da aplicação, de modo a produzir um conjunto de regiões homogêneas. Já os métodos de segmentação por detecção de bordas procuram por linhas de descontinuidades nos valores das características para gerar um mapa de bordas com regiões delimitadas por contornos.

A aquisição prévia de informações pode ser considerada como uma tarefa intrincada e até inviável em determinadas áreas de aplicação. Desta maneira, a criação de metodologias que extraíam informações a partir da própria imagem a ser segmentada passa a apresentar grande importância, entretanto, acarretando no aumento da complexidade para o desenvolvimento de um método de segmentação. Os métodos que apresentam essas características são denominados *não supervisionados*.

Dado que quaisquer informações necessárias devem ser obtidas durante o processo de segmentação, este normalmente é dividido em duas etapas. A primeira é responsável pela obtenção do número de classes contidas na imagem e a estimação dos parâmetros que as descrevem e a segunda utiliza tais parâmetros para efetuar a segmentação final.

Os métodos de segmentação por regiões comumente utilizados baseiam-se apenas em características extraídas de um conjunto de pixels. Entretanto, técnicas de segmentação que associam tais características com a dependência espacial entre as regiões da imagem têm recebido crescente atenção da comunidade científica, pois apresentam uma maior acuracidade nos resultados obtidos em relação a métodos que utilizam técnicas baseadas apenas nas características dos pixels, as quais podem gerar resultados insatisfatórios.

Uma abordagem utilizada para acrescentar a dependência espacial à segmentação de imagens é obtida por meio da formulação Bayesiana que associa as características das regiões com modelos estocásticos de imagens, esses obtidos através de campos aleatórios (Winkler 2003), como por exemplo o campo aleatório de Markov (MRF, *Markov Random Field*).

Nessa abordagem, a imagem de entrada é particionada em regiões com ou sem intersecção, a partir das quais são extraídas características, além de se utilizar um MRF para modelar a distribuição dos rótulos das classes. Finalmente, essas duas informações são associadas por meio da formulação Bayesiana e a segmentação final consiste na maximização de uma função de probabilidade ou na minimização de uma função de energia.

Dentre as principais vantagens de se utilizar a segmentação baseada em campos aleatórios estão a integração ao processo das relações espaciais entre regiões vizinhas presentes na imagem (Dubes e Jain 1989); a utilização, possibilitada pela formulação Bayesiana, de diversas características para descrição da imagem; a distribuição dos rótulos das classes para geração da segmentação final obtida diretamente do campo aleatório (Deng e Clausi 2004) e; restrições podem ser inseridas na própria função de energia a ser minimizada (Geman et al. 1990).

Devido aos bons resultados obtidos com sua utilização, as abordagens baseadas em campos aleatórios para modelagem da dependência espacial em imagens são aplicadas nas mais diversas áreas, tais como medicina (Rajapakse et al. 1997), sensoriamento remoto (Fjortoft et al. 2003), visão computacional (Geiger e Girosi 1991), recuperação de dados em bases gráficas (Farb e Jain 1996), restauração de imagens (Tonazzini e Bedini 2003), segmentação de vídeo (Luthon et al. 1999) e reconhecimento de manuscritos (Cai e Liu 2002).

Outra tarefa que apresenta alta complexidade está relacionada à definição de um conjunto de características capazes de descrever de maneira efetiva cada região contida em uma imagem. Tal descrição, normalmente é efetuada utilizando-se informações obtidas a partir de

métodos de análise de texturas (Tuceryan e Jain 1998). Entretanto, não há um único conjunto de características que possa ser utilizado de maneira satisfatória na análise de imagens em diferentes domínios de aplicações.

A utilização de cores também desempenha um papel importante na descrição de regiões contidas na imagem (Lucchese e Mitra 2001). Muitas técnicas para extração de medidas capazes de resumir as propriedades de regiões se baseiam em histogramas de cores. Embora tais técnicas sejam amplamente utilizadas em determinadas aplicações, elas apresentam a desvantagem de que as informações espaciais não são incorporadas no histograma.

Enquanto progresso significativo tem sido feito na segmentação de texturas (Tuceryan e Jain 1998) e na segmentação de imagens coloridas (Cheng et al. 2001) de maneira separada, a combinação dessas duas propriedades é considerada como um problema bem mais desafiador (Deng e Manjunath 2001). Entretanto, esta combinação deve ser explorada pois é capaz de prover informações mais acuradas para auxiliar na descrição de regiões (Palm 2004).

Considerando os aspectos descritos, este trabalho propõe um método de segmentação não supervisionado aplicável a imagens coloridas e monocromáticas. O método utiliza a formulação Bayesiana para associar a dependência espacial modelada por um campo aleatório de Markov com características de texturas. A segmentação final é obtida por meio da aplicação de técnicas de relaxação para minimizar uma função de energia definida a partir da referida associação.

1.1 Objetivos e Contribuições

Apesar do crescente interesse na abordagem de segmentação que considera a dependência espacial, há um número pequeno de trabalhos utilizando campo aleatório de Markov para análise de imagens. Dessa maneira, considera-se a escolha desta técnica de segmentação como uma contribuição para a literatura, particularmente no país, onde se tem conhecimento da existência de poucos trabalhos, tais como os propostos por Balan (2003), Bruno e Costa (2000) e Milsztajn (2003).

Faz-se uma descrição dos modelos estocásticos de imagens e das técnicas de relaxação normalmente utilizadas. Adicionalmente, os principais métodos de segmentação que consideram a dependência espacial são apresentados a partir de uma revisão bibliográfica incluindo desde os primeiros trabalhos propostos até os mais atuais.

Uma extensa revisão bibliográfica é efetuada para os métodos de análise de texturas, constituindo uma contribuição pioneira para a literatura nacional. Tais métodos, utilizados para descrição das regiões da imagem, são agrupados segundo uma pequena revisão na classificação

proposta por Tuceryan e Jain (1998).

Além das contribuições citadas, um novo método de segmentação baseada em dependência espacial é proposto. Como principais objetivos, o método visa apresentar resultados precisos e ser aplicável tanto em imagens coloridas quanto monocromáticas. Para isso, associa-se um campo aleatório de Markov com características de texturas, permitindo que a segmentação final seja obtida por meio da minimização da função de energia utilizando-se a técnica de relaxação denominada modos condicionais de iteração (ICM, *iterated conditional modes*).

Para o desenvolvimento da metodologia proposta, uma série de aspectos relevantes são discutidos, dentre eles se encontram a utilização de métodos de análise de texturas coloridas associados com dependência espacial; a estimação dos parâmetros a partir de regiões que apresentam características homogêneas; a utilização de um conjunto de critérios para determinação do número de classes presentes na imagem; a dispensabilidade de se utilizar a modelagem de texturas; a flexibilidade na definição da função de energia; e, finalmente, algumas características do algoritmo ICM.

Considerando os referidos aspectos, este trabalho apresenta as seguintes contribuições: a criação de um histograma tridimensional que auxilia na localização das classes homogêneas, a utilização de critérios, normalmente aplicados à seleção de características, para determinar o número de classes, a aplicação direta de métodos de análise de texturas a imagens coloridas para descrição das características dos pixels durante a segmentação, e a proposição de um passo adicional para o algoritmo ICM.

Finalmente, como última contribuição apresentada está uma série de experimentos efetuados visando à seleção dos métodos para análise de texturas descritos na seção 2.2. A partir dos resultados obtidos com a classificação de texturas, são analisados os principais aspectos pertinentes a este trabalho, expondo os pontos fracos e fortes apresentados pelos métodos avaliados. Aquele que apresentar o melhor desempenho será utilizado para extração de características durante a segmentação de imagens.

1.2 Organização do Trabalho

O capítulo 2 descreve os aspectos relacionados com a segmentação de imagens que utiliza a dependência espacial. Para isso, apresenta os modelos estocásticos de imagens, os métodos de análise de texturas, descreve como a formulação Bayesiana relaciona as características de texturas com a dependência espacial, e descreve os principais métodos de segmentação encontrados na literatura.

No capítulo 3 é descrita a metodologia proposta para o desenvolvimento do segmentador não supervisionado baseado em dependência espacial. Esse capítulo é dividido em três seções, a primeira descreve os aspectos a serem considerados para o desenvolvimento do método, enquanto que as duas últimas seções abordam as etapas de estimação dos parâmetros e segmentação final utilizadas no método proposto.

Vários experimentos visando à avaliação dos métodos de análise de texturas e da metodologia proposta são descritos no capítulo 4, onde as características de texturas que apresentam o melhor desempenho são utilizadas para segmentação de imagens. Finalmente, o capítulo 5 apresenta as conclusões obtidas com a elaboração deste trabalho, além de mencionar alguns trabalhos futuros a serem desenvolvidos.

CAPÍTULO 2

TRABALHOS RELACIONADOS

O método de segmentação não supervisionado proposto neste trabalho se baseia em dependência espacial, combinando características de texturas com um campo aleatório. Para isso, informações, tais como o número de classes e parâmetros que as descrevem, são extraídos a partir da imagem de entrada e a segmentação final é obtida por meio de alterações efetuadas no campo aleatório, o qual considera que os pixels localizados em uma vizinhança tendem a pertencer a uma mesma classe, ou seja, estabelece a dependência espacial entre os vizinhos.

Com intuito de estabelecer a dependência espacial entre regiões vizinhas modela-se probabilisticamente suas interações, através de um campo aleatório de Markov, que considera apenas interações locais entre os pixels, modelando a distribuição dos rótulos das classes na imagem. Os conceitos de MRF e outros modelos estocásticos para imagens são descritos na seção 2.1.

As características utilizadas, determinadas a partir da imagem de entrada, podem ser estatísticas de segunda ordem, como medidas obtidas da matriz de co-ocorrência, parâmetros de um modelo para texturas, como parâmetros de modelos estocásticos, coeficientes obtidos a partir de alguma transformada aplicada em regiões da imagem, o tom de cinza de cada pixel, ou até mesmo uma combinação dessas medidas, representadas através de um vetor de características. A seção 2.2 descreve os principais métodos de análise de texturas encontrados na literatura.

As técnicas de segmentação baseadas em dependência espacial utilizam a formulação Bayesiana para expressar a relação entre as características e o modelo estocástico empregado para descrever a distribuição dos rótulos das classes, com objetivo de maximizar a probabilidade de obtenção de uma segmentação correta para a imagem.

Com a utilização da formulação Bayesiana pode-se estimar a segmentação ótima, segundo o modelo estocástico utilizado. No entanto, a demanda computacional é elevada, dada a necessidade de calcular a probabilidade para cada uma das segmentações possíveis da imagem em questão. Por exemplo, considerando uma imagem com apenas 64×64 pixels que apresente duas classes distintas, é necessário o cálculo da probabilidade de 2^{4096} possíveis segmentações.

Com objetivo de tornar aceitável o tempo computacional são utilizadas técnicas de re-

laxação, que determinam uma segmentação aproximada da ótima. Essas técnicas são métodos iterativos que tendem a gerar segmentações com maior probabilidade de estarem corretas, sem a necessidade de determinar a probabilidade para todas as segmentações, acarretando a redução do espaço de busca (Dubes e Jain 1989). A seção 2.3 descreve os conceitos da formulação Bayesiana, bem como as principais técnicas de relaxação utilizadas para maximização da probabilidade.

Uma revisão dos principais métodos baseados em dependência espacial para segmentação de imagens coloridas e monocromáticas é apresentada na seção 2.4, tratando em maiores detalhes daqueles que dividem a segmentação em duas etapas. A primeira propicia a obtenção dos parâmetros para descrição das classes e na segunda são efetuadas alterações no MRF visando à obtenção de uma amostra que represente a segmentação final da imagem.

2.1 Modelos Estocásticos de Imagens

Após a digitalização, os pixels de uma imagem monocromática assumem valores fixos, normalmente entre 0 e 255, e o relacionamento entre cada par de pixels pode ser descrito de maneira determinística. Entretanto, considerando-se que cada pixel pode assumir valores dentro de um intervalo com uma dada probabilidade maior que zero, a imagem torna-se um campo aleatório e a interação entre seus componentes passa a ser probabilística.

Com a aleatoriedade presente na imagem, torna-se de interesse a determinação da probabilidade que cada pixel assuma um determinado valor, ou seja, a probabilidade de ocorrer uma amostra específica do campo aleatório. Para calcular essa probabilidade, considerando que o campo aleatório segue uma distribuição de probabilidade descrita por um determinado conjunto de parâmetros, em primeira análise, basta determinar a probabilidade conjunta das variáveis que compõem o campo aleatório.

Outra informação relevante a ser obtida a partir de um campo aleatório é a estimação dos parâmetros da distribuição de probabilidade responsáveis por gerar um determinado conjunto de amostras. Em outras palavras, dado um conjunto de imagens, o objetivo é determinar quais valores de parâmetros expressam a distribuição conjunta de probabilidade, onde cada pixel da imagem denota uma variável aleatória.

Os parâmetros da distribuição conjunta de um campo aleatório podem ser utilizados, por exemplo, na classificação de imagens. Durante a etapa de treinamento, estima-se os parâmetros responsáveis por gerar as texturas que descrevem cada classe. Na classificação, cada amostra é atribuída à classe que apresenta parâmetros mais próximos aos seus, segundo

alguma medida de similaridade.

Na segmentação por regiões, os parâmetros do campo aleatório podem ser utilizados em conjunto com um algoritmo de *divisão-e-união* tanto no particionamento das regiões descritas por parâmetros distintos, quanto no agrupamento de regiões que apresentam características semelhantes. Ou seja, na classificação e na segmentação de imagens, informações obtidas a partir da distribuição de probabilidade de um campo aleatório podem ser utilizadas como características de texturas.

Ainda na classificação e segmentação de imagens, os campos aleatórios são utilizados de maneira distinta daquelas descritas anteriormente. Ao invés de modelar os pixels das imagem, modela-se a distribuição dos rótulos das classes, ou seja, considerando que existam m classes distintas, os elementos do campo aleatório podem assumir valores entre 1 e m , e a amostra que maximiza a probabilidade conjunta representa a classificação ou segmentação corretas.

Embora o conjunto de parâmetros de um campo aleatório possa ser utilizado nessas aplicações, sua obtenção não se apresenta como uma tarefa simples. Alguns problemas são encontrados, tais como a falta de informações sobre a dependência entre as variáveis aleatórias, a alta complexidade computacional para determinação da probabilidade conjunta e o desconhecimento da distribuição conjunta de probabilidade seguida pelo campo aleatório.

Esta seção descreve com mais detalhes os problemas citados e algumas aproximações que possibilitam a utilização da modelagem probabilística de imagens tanto para análise de texturas quanto para obtenção de modelos para distribuição dos rótulos das classes utilizados na classificação e segmentação de imagens. Para isso, inicialmente são apresentados os conceitos de imagem aleatória (definição 2.1) e realização de uma imagem (definição 2.2).

Definição 2.1 (imagem aleatória) *Seja $\Lambda = \{0, 1, \dots, G-1\}$ o vetor aleatório $\mathbf{X} = (X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$ e cada X_i uma variável aleatória que apresenta o valor $x_i \in \Lambda$. Uma imagem aleatória é o vetor \mathbf{X} , em que $x_i \in \Lambda$ denota o valor do i -ésimo pixel enquanto G e n representam, respectivamente, o número de níveis de cinza e pixels contidos na imagem.*

Definição 2.2 (realização de uma imagem) *Chama-se de realização de uma imagem o vetor $\mathbf{x} = (x_1, x_2, \dots, x_n)$ que contém os valores das variáveis aleatórias da imagem \mathbf{X} . Como $x_i \in \Lambda$, o espaço amostral para a distribuição conjunta das variáveis pertencentes a \mathbf{X} é $\Omega = \{0, 1, \dots, G-1\}^n$, portanto $\mathbf{x} \in \Omega$.*

Com a modelagem probabilística, a imagem torna-se um conjunto de variáveis aleatórias. Pode-se então, calcular a probabilidade $P(\mathbf{x})$ de ocorrer uma realização específica, ou seja, a

probabilidade que as variáveis X_i assumam simultaneamente os valores x_i , para $i = 1, 2, \dots, n$. Tal probabilidade é utilizada na estimação dos parâmetros e na maximização da probabilidade *a posteriori* em métodos de segmentação baseados em dependência espacial.

Um dos problemas do cálculo de $P(\mathbf{x})$ se deve à falta de informações sobre a dependência entre as variáveis aleatórias (geralmente dependentes). Desta maneira, torna-se impraticável utilizar a equação 2.1 para determinação da probabilidade de uma realização específica, onde $P(X_i = x_i)$ denota a probabilidade da variável aleatória X_i assumir o valor x_i , valor também desconhecido, devido à falta de informações sobre a distribuição de probabilidade associada à variável X_i .

$$P(\mathbf{x}) = \prod_{i=1}^n P(X_i = x_i) \quad (2.1)$$

Uma possível solução para resolver o problema da dependência entre as variáveis aleatórias é proposta por Abend et al. (1965), aplicada a variáveis aleatórias binárias (ou, imagens binárias). Sugere-se um método para classificação ótima, combinando a expansão ortonormal das funções de probabilidade com a geração da distribuição conjunta de probabilidade, a partir da distribuição marginal com ordens baixas. No entanto, para a determinação de $P(\mathbf{x})$, esse método necessita calcular $2^n - 1$ funções, quantidade inviável mesmo para valores moderados de n .

Devido à alta demanda computacional para o cálculo da probabilidade conjunta $P(\mathbf{x})$, Abend et al. concluem que se deve considerar apenas a dependência local entre as variáveis aleatórias, obtendo-se assim, soluções aproximadas, entretanto computáveis. Com a utilização da dependência local são propostos diversos modelos estocásticos para imagens, tais como os presentes nos trabalhos de Abend et al. (1965), Besag (1974), Geman e Geman (1984), Kanal (1980) e Pickard (1980).

Nos trabalhos de Abend et al. (1965) e Kanal (1980) é desenvolvido o modelo estocástico de imagens denominado *Markov Mesh*. As três definições a seguir, utilizadas para descrição desse modelo, se baseiam na disposição espacial das variáveis aleatórias mostradas na figura 2.1.

$$\begin{array}{cccc} X_{1,1} & X_{1,2} & \dots & X_{1,N_2} \\ X_{2,1} & X_{2,2} & \dots & X_{2,N_2} \\ \vdots & \vdots & & \vdots \\ X_{N_1,1} & X_{N_1,2} & \dots & X_{N_1,N_2} \end{array}$$

Figura 2.1: Disposição espacial das variáveis aleatórias no modelo *Markov Mesh*.

1. $\psi_{c,d}$: matriz retangular com as variáveis aleatórias $X_{i,j}$, onde $i \leq c$ e $j \leq d$.

2. $\psi_{c,d}^{a,b}$: matriz retangular $\psi_{c,d}$ sem a variável aleatória $X_{a,b}$.
3. $Z^{a,b}$: matriz, não retangular, com as variáveis aleatórias $X_{i,j}$, $i < a$ ou $j < b$, ou seja, todas as variáveis à esquerda e acima de $X_{a,b}$.

No modelo *Markov Mesh*, a probabilidade da variável aleatória $X_{a,b}$ assumir o valor $x_{a,b}$, dadas as variáveis aleatórias que compõem $Z^{a,b}$, é mostrada pela equação 2.2, onde $U_{a,b}$ denota um subconjunto de $\psi_{a,b}^{a,b}$.

$$P(X_{a,b} = x_{a,b} | Z^{a,b}) = P(X_{a,b} = x_{a,b} | U_{a,b}) \quad (2.2)$$

Utilizando essa suposição quanto a dependência entre as variáveis, a probabilidade de ocorrência da realização \mathbf{x} pode ser obtida por meio da equação 2.3.

$$P(\mathbf{x}) = \prod_{i=1}^{N_1} \prod_{j=1}^{N_2} P(X_{i,j} = x_{i,j} | U_{i,j}) \quad (2.3)$$

A dependência de cada variável aleatória altera-se de acordo com o tamanho do conjunto $U_{a,b}$. Na figura 2.2 são mostradas algumas configurações para as dependências, os elementos pertencentes ao conjunto $U_{a,b}$ são denominados vizinhos de $X_{a,b}$. Por exemplo, utilizando-se a dependência espacial mostrada pela figura 2.2(b), a probabilidade de uma realização é dada pela equação 2.4.

$$P(\mathbf{x}) = \prod_{i=1}^{N_1} \prod_{j=1}^{N_2} P(X_{i,j} = x_{i,j} | X_{i-1,j-1}, X_{i-1,j}, X_{i,j-1}). \quad (2.4)$$

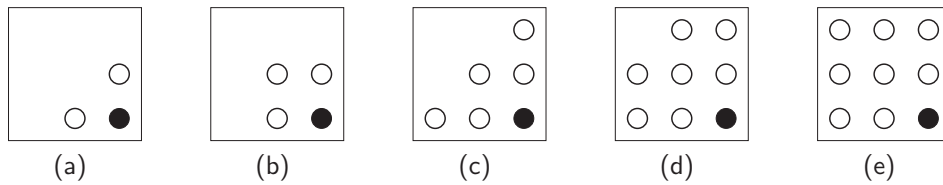


Figura 2.2: Configurações para dependência local no modelo *Markov Mesh*, o círculo preenchido indica a variável aleatória $X_{a,b}$.

Uma abordagem semelhante à proposta por Abend et al. é desenvolvida no trabalho de Bartlett (1968), que considera uma vizinhança não somente à esquerda superior da variável aleatória, mas também na região inferior direita. Um exemplo para o cálculo da probabilidade de uma realização é mostrada na equação 2.5. No entanto, alguns problemas são identificados por Besag (1972), como a inexistência de um método direto para determinação da distribuição conjunta de probabilidade e a existência de condições de consistência altamente restritivas.

$$P(\mathbf{x}) = \prod_{i=1}^{N_1} \prod_{j=1}^{N_2} P(X_{i,j} = x_{i,j} | X_{i-1,j}, X_{i+1,j}, X_{i,j-1}, X_{i,j+1}). \quad (2.5)$$

Besag (1972) apresenta uma solução consistente para o cálculo das probabilidades condicionais mostradas pela equação 2.5, no entanto, esta solução é específica para considerações de que o valor do pixel central depende apenas dos valores de seus quatro vizinhos. Já em 1974, Besag generaliza para vizinhança de tamanho arbitrário sem perder a consistência, propondo que a modelagem da dependência local entre as variáveis seja efetuada por meio de um campo aleatório de Markov.

Estabelecendo um relacionamento entre o campo aleatório de Markov e o campo aleatório de Gibbs (GRF, *Gibbs Random Field*), Geman e Geman (1984) apresentam resultados que possibilitam o cálculo de $P(\mathbf{x})$ e $P(X_s = x_s | X_t = x_t, s \neq t)$, mostrado na definição 2.7. As seguintes definições são utilizadas para enunciar o teorema 2.1, que relaciona MRF com um subconjunto de GRF.

Definição 2.3 (campo aleatório) *Seja $X = \{X_1, X_2, \dots, X_n\}$ um conjunto de variáveis aleatórias, $\mathbf{x} = (x_1, x_2, \dots, x_n)$ uma realização pertencente ao espaço amostral $\Omega = \{0, 1, \dots, G-1\}^n$ e P uma medida de probabilidade em Ω . P é denominado campo aleatório em Ω , se $P(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in \Omega$.*

Definição 2.4 (sistema de vizinhança) *Seja $S = \{1, 2, \dots, n\}$. Uma coleção $\partial = \{\partial_s : s \in S\}$ de conjuntos é chamada sistema de vizinhança, se $s \notin \partial_s$ e $s \in \partial_t$ se, e somente se, $t \in \partial_s$. Os elementos $t \in \partial_s$ são denotados vizinhos de s .*

Definição 2.5 (ordem da vizinhança) *A variável aleatória localizada na posição (i, j) tem como vizinhos as variáveis com coordenadas (k, l) , tal que $0 < (k - i)^2 + (l - j)^2 \leq c$, onde c denota a ordem da vizinhança.*

Definição 2.6 (clique) *O subconjunto C de S é um clique se para quaisquer dois elementos distintos $s, t \in C$, $s \in \partial_t$. O conjunto composto pelos cliques é denotado por ϕ .*

A definição de sistema de vizinhança será utilizada para estabelecer a dependência local entre as variáveis aleatórias, pois uma característica do GRF é a determinação de $P(\mathbf{x})$, $\mathbf{x} \in \Omega$, através das relações locais entre seus componentes. A figura 2.3 mostra configurações para vizinhanças de primeira ($c = 1$) e segunda ordem ($c = 2$) e seus respectivos cliques.

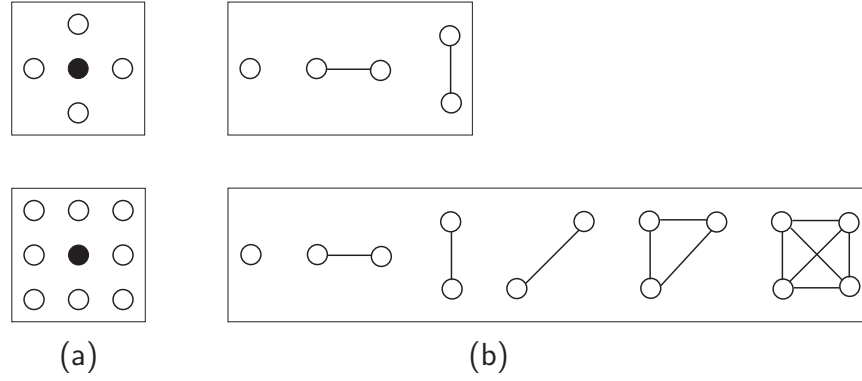


Figura 2.3: Vizinhança em relação ao pixel central e cliques. (a) vizinhança de primeira ordem e segunda ordem, respectivamente; (b) cliques das respectivas vizinhanças.

Definição 2.7 (características locais) *As características locais de uma medida de probabilidade P definida em Ω é a probabilidade condicional da forma*

$$P(X_s = x_s | X_t = x_t, s \neq t), \quad (2.6)$$

ou seja, a probabilidade da variável aleatória X_s assumir o valor x_s dadas todas outras variáveis.

Definição 2.8 (campo aleatório de Markov) *O campo aleatório P é um MRF com respeito ao sistema de vizinhança ∂ se para todo $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \Omega$*

$$P(X_s = x_s | X_t = x_t, t \neq s) = P(X_s = x_s | X_t = x_t, t \in \partial_s), \quad (2.7)$$

isto é, as características locais dependem apenas da vizinhança.

Alguns trabalhos, como o desenvolvido por Cross e Jain (1983), exigem que P apresente a propriedade de homogeneidade, isto é, a probabilidade de uma variável aleatória depende somente da configuração dos vizinhos, sendo invariante com respeito à translação. A propriedade de homogeneidade, por exemplo, proporciona significativa redução do número de parâmetros necessários para a modelagem de texturas.

A motivação para relacionar MRF e campo aleatório de Gibbs vem da dificuldade do cálculo das características locais utilizando somente MRF, por desconhecer a distribuição conjunta das variáveis aleatórias, como mostra a equação 2.8. Onde \mathbf{x}^x representa a realização em que $X_s = x$ e as demais variáveis aleatórias apresentam os mesmos valores que em \mathbf{x} .

$$P(X_s = x | X_t = x_t, s \neq t) = \frac{P(\mathbf{x})}{\sum_{\mathbf{x} \in \Lambda} P(\mathbf{x}^x)} \quad (2.8)$$

O estabelecimento da relação entre os dois campos aleatórios possibilita o cálculo da probabilidade de uma realização em Ω e o cálculo das características locais de um MRF utilizando-se os resultados obtidos para o campo aleatório de Gibbs.

A medida de probabilidade mostrada pela equação 2.9 é denominada campo aleatório de Gibbs induzido pela função de energia H , e Z é denotada função de partição. Dado um sistema de vizinhança, a função de energia representa a intensidade das relações entre as variáveis aleatórias. Cada contribuição para essa função é dada pelas funções de potencial, mostradas na definição 2.9, tais funções dependem apenas das variáveis aleatórias dentro de uma vizinhança.

$$P(\mathbf{x}) = \frac{1}{Z} \exp(-H(\mathbf{x})), \quad Z = \sum_{\mathbf{z} \in \Omega} \exp(-H(\mathbf{z})) \quad (2.9)$$

Definição 2.9 (funções de potencial) *Seja $S = \{1, 2, \dots, n\}$. Denota-se potencial, a família $\{U_A : A \subset S\}$ de funções em Ω tal que*

1. $U_\emptyset = 0$
2. $U_A(\mathbf{x}) = U_A(\mathbf{y})$ se $x_s = y_s$ para cada $s \in A$.

A energia de U é dada por $H_U = \sum_{A \subset S} U_A$, denotada como função de energia em relação ao potencial U . U é chamada de potencial de vizinhança com respeito ao sistema de vizinhança ∂ , se A não é um clique então $U_A = 0$, ou seja, depende apenas da vizinhança.

Definição 2.10 (campo aleatório de Gibbs) *Um campo aleatório P é um campo aleatório de Gibbs com relação ao potencial U se apresenta a forma*

$$P(\mathbf{x}) = \frac{1}{Z} \exp(-H_U(\mathbf{x})). \quad (2.10)$$

Se U é um potencial de vizinhança então P é chamado campo aleatório de Gibbs de vizinhança (NGRF, Neighbor Gibbs Random Field).

Teorema 2.1 *Dado o sistema de vizinhança ∂ em S , um campo aleatório é um MRF em ∂ se, e somente se, ele é um NGRF em ∂ .*

A prova do teorema 2.1 encontra-se nos trabalhos de Besag (1974), Kindermann e Snell (1980) e Winkler (2003). O resultado obtido permite que a probabilidade de uma realização de um MRF seja calculada por meio da equação 2.10. No entanto, o esforço computacional

apresenta-se extremamente alto, pois o cálculo do denominador depende da determinação das funções de potencial para todas as todas as realizações possíveis (Geman e Geman 1984).

As características locais em um NGRF podem ser facilmente calculadas devido ao resultado mostrado pela equação 2.11 (Geman e Geman 1984). Onde \mathbf{x}^x denota a realização em que a variável aleatória $X_s = x$ e as demais variáveis apresentam os mesmos valores que na realização \mathbf{x} .

$$P(X_s = x_s | X_t = x_t, s \neq t) = \frac{\exp\left(-\sum_{C:s \in \phi} U_C(\mathbf{x})\right)}{\sum_{\mathbf{x} \in \Lambda} \exp\left(-\sum_{C:s \in \phi} U_C(\mathbf{x}^x)\right)} \quad (2.11)$$

Alguns modelos para o cálculo de características locais são especificados, entre eles, o de Ising (citado por Geman e Geman (1984)) e o de Potts (citado por Rosholm (1997)), ambos provêm da física estatística e foram criados a partir do campo aleatório de Gibbs.

O modelo proposto por Ising em 1925 teve como objetivo modelar interações eletromagnéticas em compostos metálicos. Este modelo considera apenas dois estados possíveis, portanto, gerando o espaço amostral $\Omega = \{0, 1\}^n$. A probabilidade de uma realização é dada pela equação 2.12, e as características locais são mostradas na equação 2.13, onde $i \sim j$ denota que as variáveis i e j formam um clique dentro da vizinhança, n_1 denota o número de vizinhos de X_i que apresentam o valor 1 e α, β são os parâmetros da distribuição.

$$P(\mathbf{x}) = \frac{\exp\left(-\alpha \sum_i x_i - \beta \sum_{i \sim j} x_i x_j\right)}{\sum_{\mathbf{y} \in \Omega} \exp\left(-\alpha \sum_i y_i - \beta \sum_{i \sim j} y_i y_j\right)} \quad (2.12)$$

$$P(X_s = x_s | \partial_s) = \frac{\exp(-x_s(\alpha + \beta n_1))}{1 + \exp(-\alpha - \beta n_1)} \quad (2.13)$$

Percebe-se que a equação 2.12 pode ser obtida com a utilização da definição 2.10, fazendo o potencial U da forma $(\alpha x_i + \beta x_i x_j)$, isso porque o modelo de Ising apresentado utiliza apenas cliques compostos por no máximo dois elementos, onde α e β são parâmetros que representam as forças magnéticas entre os vizinhos (Kindermann e Snell 1980).

Em 1952, Potts propõe um modelo que generaliza o criado por Ising, permitindo a observação de fenômenos que apresentam mais de dois estados. Considerando G estados possíveis, o espaço amostral do modelo de Ising torna-se $\Omega = \{0, 1, \dots, G - 1\}^n$, para

n variáveis aleatórias. A probabilidade de uma realização é a mesma do modelo de Ising (equação 2.12), as características locais podem ser obtidas pela equação 2.14, onde α_k , α_l e β_k são os parâmetros da distribuição e $n_i(k)$ denota o número de vizinhos de X_i que apresentam o valor k .

$$P(X_s = k | \partial_s) = \frac{\exp(-\alpha_k - \beta_k n_i(k))}{\sum_{l=0}^{G-1} \exp(-\alpha_l - \beta_k n_i(l))} \quad (2.14)$$

2.2 Características de Texturas

Uma das tarefas mais complexas presente na análise de imagens está em definir um conjunto de características capazes de descrever de maneira efetiva cada região contida em uma imagem de modo a ser utilizado em processos de mais alto nível, tais como o reconhecimento e análise. Para isso, uma abordagem natural está em recorrer às características utilizadas pelos humanos na interpretação de informações visuais.

A textura encontra-se entre as características empregadas pelo sistema visual humano, contendo informações sobre a distribuição espacial e a variação de luminosidade, além de descrever o arranjo estrutural das superfícies e relações entre regiões vizinhas. Dessa maneira, a utilização de informações texturais se apresenta como uma abordagem adequada para descrição de regiões da imagem.

Embora o sistema visual humano apresente facilidade no reconhecimento e descrição de texturas, é extremamente difícil formalizar sua definição ou desenvolver um conjunto de descritores que possam ser utilizados para análise de imagens em diferentes domínios de aplicações. Tal dificuldade é refletida pela grande quantidade de definições e métodos de análise de texturas encontrados na literatura (Randen e Husoy 1999, Reed e Dubuf 1993, Zhang e Tan 2002).

Exemplos de definições para textura podem ser encontrados no trabalho de Tamura et al. (1978), que a define como constituinte de uma região macroscópica, onde sua estrutura é devido a repetição de padrões, nos quais seus elementos ou primitivas são arranjados conforme uma regra de composição. Rosenfeld e Troy (1970) definem textura como um arranjo repetitivo de padrões sobre uma área e tentam medir sua aspereza observando fatores como a dependência de níveis de cinza e autocorrelação.

Segundo Haralick (1979), uma textura pode ser descrita pela interação entre as primitivas tonais que a compõe, estas ocorrendo em diferente número e formas. Pixels contíguos que apresentam propriedades semelhantes formam cada uma das primitivas, dentre as quais podem

ocorrer interações aleatórias ou com certo grau de dependência. Dessa maneira, a ocorrência de interações aleatórias e grandes variações no nível de cinza entre as primitivas caracterizam *texturas finas*, enquanto interações melhor definidas e a presença de regiões mais homogêneas caracterizam *texturas ásperas*.

Dentre as principais aplicações que utilizam análise de texturas estão a *classificação*, que objetiva a criação de uma mapa onde cada região homogênea é identificada como pertencente a uma determinada classe; a *segmentação*, visando à determinação das bordas entre as diferentes regiões texturizadas contidas em uma imagem e; finalmente, a *síntese de texturas*, responsável pela determinação de um modelo capaz de gerar uma dada textura. A figura 2.4 exemplifica as aplicações da análise de texturas.

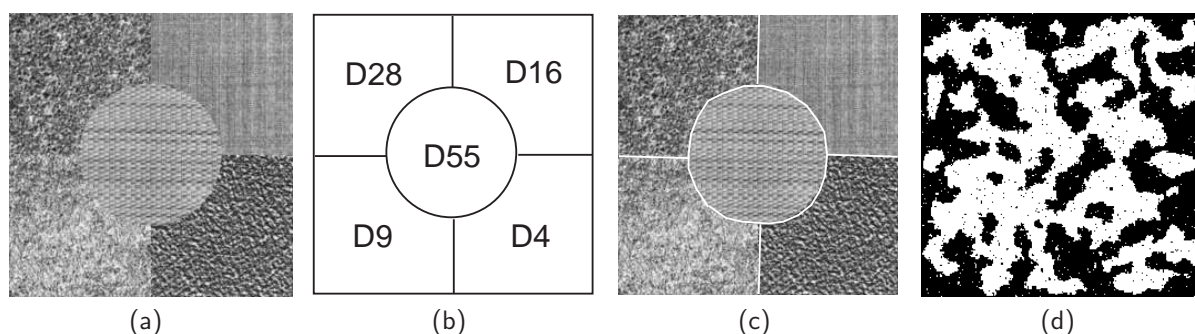


Figura 2.4: Principais aplicações da análise de texturas. (a) imagem monocromática original; (b) mapa de classes resultante da classificação da imagem, indicando as texturas presentes; (c) bordas identificadas pelo processo de segmentação; (d) textura sintética gerada pela amostragem de um campo aleatório de Markov.

Normalmente, os métodos que efetuam análise de texturas são obtidos através dos processos de extração e seleção de características. A extração é responsável por executar transformações nos dados de entrada, de modo a descrevê-los de maneira representativa e simplificada, e a seleção visa reduzir o número de características, bem como eliminar aquelas que apresentem redundância.

O restante desta seção descreve os principais métodos de extração de características de texturas. Embora não haja um consenso na literatura, neste trabalho os métodos são divididos entre as abordagens *estatística* (seção 2.2.1), *baseada em processamento de sinais* (seção 2.2.2), *geométrica* (seção 2.2.3) e *baseada em modelos paramétricos* (seção 2.2.4). Essa classificação baseia-se naquela proposta por Tuceryan e Jain (1998), entretanto, com a especificação de quais modelos são utilizados, pois entende-se que o termo *modelo* utilizado por Tuceryan e Jain, possui significado bastante amplo.

2.2.1 Abordagem Estatística

A primitiva mais simples que pode ser definida em uma imagem digital em níveis de cinza é um pixel, que tem como propriedade o próprio nível de cinza. Conseqüentemente, a distribuição dos níveis de cinza dos pixels pode ser descrita por estatísticas de primeira ordem, como média, variância, desvio padrão, obtidos por meio de um histograma computado a partir dessa distribuição.

Como as estatísticas de primeira ordem dependem apenas do nível de cinza individual de cada pixel, estas tornam-se sensíveis, por exemplo, quanto às variações do histograma. Para evitar esse tipo de problema são utilizadas estatísticas de segunda ordem, as quais dependem das transições dos níveis de cinza entre dois pixels. No entanto, ainda resultando em problemas caso haja alterações nas relações espaciais entre os pixels contidos na imagem.

Os métodos que utilizam a abordagem estatística não buscam compreender explicitamente a estrutura hierárquica da textura, mas sim tentam representar a textura indiretamente por propriedades não determinísticas que definem distribuições e relacionamentos entre os níveis de cinza dos pixels de uma imagem. Dentre os principais métodos, estão os baseados em estatísticas de primeira ordem, na matriz de co-ocorrência e em características obtidas a partir de estatísticas de ordem superior, como por exemplo a matriz de comprimento de corridas de cinza. Esses métodos são descritos a seguir.

Estatísticas de Primeira Ordem

O histograma dos níveis de cinza provê informações estatísticas de primeira ordem, pois cada pixel é considerado de maneira individual durante a determinação da distribuição dos níveis de cinza. Seja uma imagem composta por n pixels, a função de massa de probabilidade pode ser determinada por meio da equação 2.15, onde $h(i)$ denota o número de ocorrências do i -ésimo tom de cinza.

$$P(i) = \frac{h(i)}{n} \quad (2.15)$$

Embora as estatísticas de primeira ordem apresentem algumas desvantagens, o custo computacional apresentado na determinação das características é baixo pois utilizam-se apenas medidas simples, tais como média, variância, assimetria e curtose. A média, mostrada na equação 2.16, representa o valor esperado da distribuição, enquanto a variância, definida na equação 2.17, descreve quanto os valores estão dispersos em torno da média, onde H_g denota o maior nível de cinza presente na imagem.

$$\mu = \frac{1}{n} \sum_{i=0}^{H_g} h(i) \quad (2.16)$$

$$\sigma^2 = \frac{1}{n-1} \sum_{i=0}^{H_g} (h(i) - \mu)^2 \quad (2.17)$$

O grau de assimetria da distribuição, definido na equação 2.18, é um indicador da concentração de valores em relação a sua mediana, ou seja, distribuições como a Gaussiana apresentam assimetria nula, enquanto assimetria negativa está presente em distribuições que se concentram à esquerda da mediana. A curtose, definida na equação 2.19, indica o achatamento da função de distribuição, apresentando valores negativos em distribuições que apresentam forma mais achatada que a Gaussiana. A figura 2.5 ilustra os valores assumidos pelas medidas de assimetria e curtose.

$$s = \frac{1}{n\sigma^3} \sum_{i=0}^{H_g} (h(i) - \mu)^3 \quad (2.18)$$

$$k = \frac{1}{n\sigma^4} \sum_{i=0}^{H_g} (h(i) - \mu)^4 - 3 \quad (2.19)$$

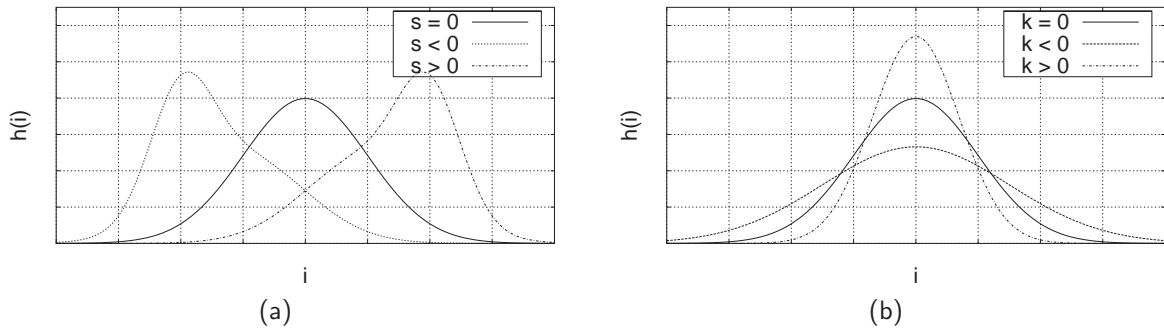


Figura 2.5: Medidas calculadas a partir do histograma. (a) coeficiente de assimetria; (b) curtose.

Outras duas medidas a serem calculadas a partir do histograma são a energia (equação 2.20) e a entropia (equação 2.21). Além das medidas apresentadas, há outras que podem ser utilizadas como características de texturas a serem extraídas a partir do histograma, tais como valor mínimo, valor máximo, moda e mediana.

$$E = \sum_{i=0}^{H_g} (P(i))^2 \quad (2.20)$$

$$H = - \sum_{i=0}^{H_g} P(i) \log(P(i)) \quad (2.21)$$

Matriz de Co-ocorrência

Métodos baseados em estatísticas de segunda ordem têm demonstrado alto poder discriminativo em uma grande variedade de imagens. Uma das abordagens utilizadas para adquirir informações sobre transições de níveis de cinza entre dois pixels é aquela obtida através da construção da matriz de co-ocorrência, baseada na ocorrência repetida da configuração de alguns níveis de cinza na textura, variando rapidamente segundo a frequência espacial em texturas finas e lentamente em texturas ásperas.

Dada uma relação espacial entre os pixels componentes de uma textura, os elementos da matriz de co-ocorrência descrevem a frequência com que ocorrem as transições de nível de cinza entre pares de pixels. Dessa maneira, efetuando-se variações na relação espacial, por meio de alterações na orientação e distância entre as coordenadas dos pixels, podem ser obtidas diversas matrizes de co-ocorrência, a partir das quais são extraídas medidas utilizadas para análise da texturas.

O número de linhas e colunas dessa matriz é proporcional à quantidade de níveis de cinza contidos na textura, ou seja, independe das dimensões da textura, acarretando, desta maneira, perda do relacionamento espacial nela contida. Adicionalmente, o elemento $P(m, n)$ da matriz de co-ocorrência representa o número de transições entre os níveis de cinza m e n que ocorrem na textura.

Entretanto, antes que essa matriz seja computada, deve-se definir as relações espaciais entre os pixels, ou seja, determinar quais pixels e quais relacionamentos espaciais serão considerados. Para isso, constrói-se o conjunto S , no qual cada elemento é composto de dois pares ordenados denotando as coordenadas de cada pixel envolvido na relação espacial. Após a determinação desse conjunto, utiliza-se a equação 2.22 para determinar o número de transições que ocorrem entre cada par de tons de cinza contido na textura, onde $f(x, y)$ denota o tom de cinza do pixel localizado na coordenada (x, y) , ou seja, na x -ésima coluna e y -ésima linha.

$$P(m, n) = \#\{(i, j), (k, l) \in S \mid f(i, j) = m \text{ e } f(k, l) = n\} \quad (2.22)$$

Uma vez determinado o número de ocorrências de cada uma das transições de níveis de cinza, basta acrescentar o valor $P(m, n)$ na m -ésima linha e n -ésima coluna da matriz, obtendo-se então a matriz de co-ocorrência. Percebe-se claramente que as dimensões de tal matriz dependem do número de níveis de cinza contidos na textura. A figura 2.6 exemplifica a composição da matriz para 4 níveis de cinza.

	0	1	2	3
0	$P(0, 0)$	$P(0, 1)$	$P(0, 2)$	$P(0, 3)$
1	$P(1, 0)$	$P(1, 1)$	$P(1, 2)$	$P(1, 3)$
2	$P(2, 0)$	$P(2, 1)$	$P(2, 2)$	$P(2, 3)$
3	$P(3, 0)$	$P(3, 1)$	$P(3, 2)$	$P(3, 3)$

Figura 2.6: Composição da matriz de co-ocorrência. Cada elemento é composto pelo número de transições específicas entre os níveis de cinza.

Apesar da matriz de co-ocorrência ter sido definida como sendo composta pelo número de transições que ocorrem entre determinados níveis de cinza, as características de texturas são obtidas a partir de outra representação desta matriz, denominada representação normalizada. Nessa representação, conforme mostrado pela equação 2.23, cada elemento da matriz original é dividido pela soma de todos os seus componentes, onde H_g denota o maior nível de cinza presente na textura.

$$p_{m,n} = \frac{P(m,n)}{\sum_{i=0}^{H_g} \sum_{j=0}^{H_g} P(i,j)} \quad m, n = 0, \dots, H_g \quad (2.23)$$

Como apresentado na equação 2.22, a construção da matriz de co-ocorrência depende das transições dos níveis de cinza entre os componentes do conjunto S . Dessa maneira, pode-se definir arbitrariamente a distância e o ângulo entre os pixels onde serão computadas as transições apenas efetuando alterações nesse conjunto, entretanto, distâncias e ângulos distintos acabam sendo incluídos em uma mesma matriz.

Visando à separação das transições para diferentes ângulos e distâncias, Haralick et al. (1973) definem de modo mais específico quais transições devem ser utilizadas para criação de cada matriz de co-ocorrência. Dois parâmetros adicionais são acrescentados na definição de P , são eles d e θ , exercendo controle sobre a distância e o ângulo entre os pixels, respectivamente. Desta maneira, diversas matrizes podem ser criadas com pequenas alterações nos parâmetros, proporcionando a obtenção de um maior número de características de texturas.

Ainda no trabalho de Haralick et al. são definidos quatro ângulos para substituir o parâmetro θ , os quais indicam como deve ser o relacionamento entre dois pixels, conforme mostra a

figura 2.7. Considerando um d fixo, para cada um desses ângulos será computada uma matriz de co-ocorrência, que representa as transições de níveis de cinza entre pixels dispostos nessa orientação específica. Por exemplo, tomando como referência o pixel central da figura 2.7, a matriz que considera a orientação de 45° baseia-se apenas nas transições entre os pixels 3, 6 e o central.

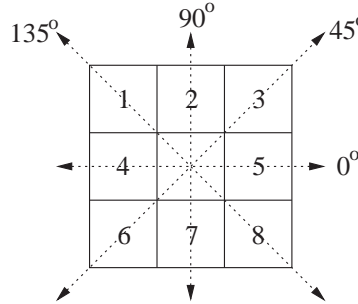


Figura 2.7: Ângulos utilizados para o cálculo das matrizes de co-ocorrência. As transições para cada ângulo são computadas sempre em relação ao pixel localizado na posição central.

Com a especificação desses quatro ângulos para expressar o relacionamento espacial, a definição de matriz de co-ocorrência apresentada anteriormente pode ser estendida pelas equações 2.24 a 2.27, onde o conjunto S é composto por todos os pares de pixels da imagem. Por exemplo, para calcular as transições que apresentam ângulo zero, deve-se relacionar o pixel central com seus vizinhos da direita e esquerda, segundo o valor escolhido para d .

$$P(i, j, d, 0^\circ) = \#\{(k, l), (m, n)\} \in S \mid |k - m| = d, l - n = 0\}, \quad (2.24)$$

$$f(k, l) = i, f(m, n) = j\}$$

$$P(i, j, d, 45^\circ) = \#\{(k, l), (m, n)\} \in S \mid (k - m = -d, l - n = d) \quad (2.25)$$

$$\text{ou } (k - m = d, l - n = -d), f(k, l) = i, f(m, n) = j\}$$

$$P(i, j, d, 90^\circ) = \#\{(k, l), (m, n)\} \in S \mid k - m = 0, |l - n| = d\}, \quad (2.26)$$

$$f(k, l) = i, f(m, n) = j\}$$

$$P(i, j, d, 135^\circ) = \#\{(k, l), (m, n)\} \in S \mid (k - m = d, l - n = d) \quad (2.27)$$

$$\text{ou } (k - m = -d, l - n = -d), f(k, l) = i, f(m, n) = j\}$$

Com objetivo de descrever as propriedades contidas em texturas, Haralick et al. (1973) propõem 14 medidas estatísticas a serem calculadas a partir das matrizes de co-ocorrência, embora apenas seis delas sejam mais relevantes, são elas: *segundo momento angular*, *entropia*, *contraste*, *variância*, *correlação* e a *homogeneidade* (Baraldi e Parmiggiani 1995).

Haralick et al. apresentam a formulação sem esclarecê-la, desta maneira, a descrição aqui apresentada se baseia no trabalho de Baraldi e Parmiggiani.

O segundo momento angular, também conhecido por energia ou uniformidade, expressa a uniformidade de uma textura. Em texturas ásperas, poucos elementos da matriz de co-ocorrência normalizada apresentam valores diferentes de zero, e quando ocorrem, são próximos de um, neste caso f_{sma} , mostrado na equação 2.28, apresenta valores próximos de um, o máximo possível para esta medida.

$$f_{sma} = \sum_{i=0}^{H_g} \sum_{j=0}^{H_g} p_{i,j}^2 \quad (2.28)$$

A medida de entropia mede a desordem contida na textura. Quando uma imagem não é uniforme, as entradas $p_{m,n}$ apresentam valores próximos de zero e f_{ent} , mostrada na equação 2.29, resulta em valores altos, entretanto não normalizados entre zero e um como ocorre na medida de segundo momento angular.

$$f_{ent} = - \sum_{i=0}^{H_g} \sum_{j=0}^{H_g} p_{i,j} \log(p_{i,j}) \quad (2.29)$$

O contraste caracteriza-se pela diferença entre os tons de cinza. Baixo contraste ocorre quando há pequena diferença entre os níveis de cinza dos pixels localizados em uma região contígua da imagem. Em termos da matriz de co-ocorrência, valores baixos para medida f_{con} , mostrada na equação 2.30, ocorrem quando há grande concentração de elementos significativos em torno da diagonal principal.

$$f_{con} = \sum_{n=0}^{H_g} n^2 \left\{ \sum_{i=0}^{H_g} \sum_{j=0}^{N_g} p_{i,j} \right\} \quad |i - j| = n \quad (2.30)$$

Medida de heterogeneidade, a variância apresenta valores altos quando os tons de cinza desviam do nível de cinza médio, independentemente da localização dos elementos da matriz e da frequência espacial que a textura apresenta. A equação 2.31 apresenta a variância, onde μ denota o valor médio da matriz de co-ocorrência.

$$f_{var} = \sum_{i=0}^{H_g} \sum_{j=0}^{H_g} (i - \mu)^2 p_{i,j} \quad (2.31)$$

A correlação mede a dependência linear entre os tons de cinza presentes em uma imagem. Valores altos de f_{corr} (equação 2.32), próximos de um, indicam a existência de relação linear

entre os pares de níveis de cinza. Nessa equação, σ_x e σ_y indicam o desvio padrão das distribuições marginais, assim como μ_x e μ_y indicam o valor médio dessas distribuições.

$$f_{\text{corr}} = \frac{1}{\sigma_x \sigma_y} \sum_{i=0}^{H_g} \sum_{j=0}^{H_g} ij p_{i,j} - \mu_x \mu_y \quad (2.32)$$

A última medida representativa proposta por Haralick et al. é a homogeneidade, mostrada na equação 2.33, a qual assume valores altos quando a textura apresenta pequenas variações de níveis de cinza entre pares de pixels. Tal medida apresenta correlação inversa com a medida de contraste mostrada na equação 2.30.

$$f_{\text{hom}} = \sum_{i=0}^{H_g} \sum_{j=0}^{H_g} \frac{1}{1 + (i - j)^2} p_{i,j} \quad (2.33)$$

Algumas variações para matriz de co-ocorrência têm sido propostas, entre elas está a matriz de co-ocorrência tridimensional, proposta por Dacheng et al. (2002), utilizada para descrição de texturas coloridas que estejam no espaço cromático HSI, quantizado em 8, 4 e 4 tons para os canais H, S e I, respectivamente.

A idéia central do método desenvolvido por Dacheng et al. é relacionar tanto os pixels vizinhos, como efetuado pela abordagem tradicional de matriz de co-ocorrência, quanto extrair o relacionamento existente entre as bandas do HSI, e com isso, gerar matrizes de co-ocorrência tridimensionais que resumem essas transições. A figura 2.8 ilustra como são consideradas as variações de intensidades em cada canal de cor.

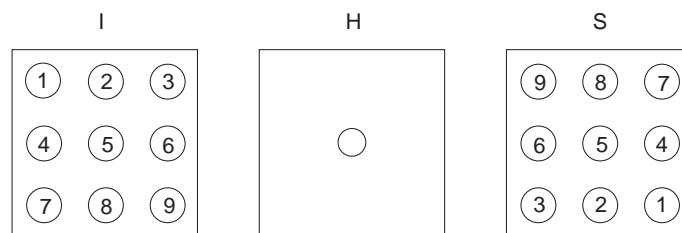


Figura 2.8: Relações entre os canais de cores, onde os números iguais indicam as variações de intensidades consideradas, sempre passando pelo pixel central do canal H.

Conforme mostrado na figura 2.8, nove direções são consideradas para descrever as relações inter-canais e inter-bandas, com isso são criadas nove matrizes, uma para extrair as características de cada relacionamento. Dada a quantização apresentada anteriormente, essas matrizes apresentam dimensões $8 \times 4 \times 4$ e o elemento (h, s, i) contém o número de transições que apresentam valor h , s e i , nas bandas H, S e I, enquanto $p_{h,s,i}$ denota a probabilidade de ocorrer tal transição.

Dentre as características a serem extraídas dessas matrizes, estão quatro medidas calculadas para matriz de co-ocorrência bidimensional, entretanto, estendidas para três dimensões. Sendo elas o segundo momento angular (equação 2.34), contraste (equação 2.35), correlação (equação 2.36) e a entropia (equação 2.37), onde G_h , G_s , G_i denotam o valor do tom máximo em cada um dos canais do espaço cromático HSI. Dessa maneira, a partir das nove matrizes são obtidas 36 características.

$$f_{asm} = \sum_{h=0}^{G_h} \sum_{s=0}^{G_s} \sum_{i=0}^{G_i} p_{h,s,i}^2 \quad (2.34)$$

$$f_{con} = \frac{1}{G_s} \sum_{k=0}^{G_s} \sum_{n=0}^{\max(G_h, G_i)} n^2 \sum_{|i-j|=n} p_{j,k,i} +$$

$$\frac{1}{G_h} \sum_{j=0}^{G_h} \sum_{n=0}^{\max(G_s, G_i)} n^2 \sum_{|i-k|=n} p_{j,k,i} +$$

$$\frac{1}{G_i} \sum_{i=0}^{G_i} \sum_{n=0}^{\max(G_s, G_h)} n^2 \sum_{|k-j|=n} p_{j,k,i} \quad (2.35)$$

$$f_{corr} = \sum_{h=0}^{G_h} \sum_{s=0}^{G_s} \sum_{i=0}^{G_i} \frac{hsip_{h,s,i} - \mu_h \mu_s \mu_i}{\sigma_h \sigma_s \sigma_i} \quad (2.36)$$

$$f_{ent} = - \sum_{h=0}^{G_h} \sum_{s=0}^{G_s} \sum_{i=0}^{G_i} p_{h,s,i} \log(p_{h,s,i}) \quad (2.37)$$

Idéia semelhante à matriz de co-ocorrência utilizada para descrever imagens coloridas proposta por Dacheng et al. é abordada no trabalho de Palm (2004). Em seu trabalho, Palm utiliza o espaço cromático LUV e o relacionamento entre os canais de cores é efetuado apenas dois a dois, ao invés de considerar os três canais de maneira simultânea.

Embora seja amplamente utilizada em análise de texturas, a matriz de co-ocorrência apresenta alguns problemas, como a inexistência de um método para seleção da distância d entre os pixels, embora possa ser reduzido por meio do cálculo de matrizes para valores distintos de d , o que acarreta no aumento do custo computacional. Além do fato dessas matrizes não capturarem os aspectos da forma das primitivas contidas na textura.

Matrizes de Comprimento de Corridas de Cinza

A amostragem de regiões colineares que compõem uma imagem permite que sejam encontradas corridas de cinza, compostas por pixels consecutivos que apresentam os mesmos

valores de nível de cinza. Com o objetivo de sintetizar as informações obtidas a partir dessas corridas, Galloway (1975) propõe a criação de matrizes cujos elementos contêm o número de corridas com um dado tamanho para um determinado nível de cinza, provendo desta maneira, estatísticas de ordem superior a serem utilizadas na análise de texturas.

A partir dessas matrizes, denominadas matrizes de comprimento de corridas de cinza (GLRLM, *Gray Level Run Length Matrices*), podem ser obtidas informações relevantes sobre as características da textura que está sendo analisada. Em uma textura áspera espera-se que corridas relativamente longas sejam freqüentes, enquanto em texturas finas ocorram corridas mais curtas, devido a presença de bordas, responsáveis por variações rápidas e acentuadas nos níveis de cinza dos pixels pertencentes a uma determinada linha da imagem, causando desta maneira, interrupções freqüentes nas corridas.

Para que as matrizes propostas por Galloway sejam criadas, define-se que um conjunto composto por pixels consecutivos, apresentando o mesmo nível de cinza, colineares em uma dada orientação representa uma *corrida de cinza* e o número de pixels contidos neste conjunto denota o *tamanho da corrida*. Desta maneira, cada elemento da matriz, representado por $P(i, j|\theta)$, contém o número de corridas com tamanho j , tendo i como o nível de cinza de seus pixels, onde o parâmetro θ indica a orientação do segmento de reta formado pelos pixels.

Devido ao grande número de níveis de cinza presente nas imagens e a possível presença de ruído, essas matrizes podem ser construídas utilizando agrupamentos de tons de cinza consecutivos, ou seja, cada i de $P(i, j|\theta)$ pode representar um intervalo de níveis de cinza. Por exemplo, Weszka et al. (1976) utilizam os intervalos lineares $(0, 7)$, $(8, 15)$, ..., $(56, 63)$ para imagens compostas por 64 níveis de cinza.

Normalmente, a matriz GLRL contém elementos que apresentam valores baixos nas colunas localizadas mais a direita, ou seja, aquelas que representam as corridas de comprimento longo. Para evitar que as matrizes fiquem esparsas, os tamanhos das corridas são particionados logaritmicamente em intervalos, como por exemplo, a partição 1, 2-3, 4-7, 8-15, 16-31 para corridas com tamanho entre 1 e 31 (Albregtsen et al. 2000).

Para uma dada imagem, pode-se calcular um conjunto de matrizes GLRL para corridas em quaisquer orientações. No entanto, assim como na criação das matrizes de co-ocorrência, geralmente são calculadas apenas as matrizes para corridas nas direções 0° , 45° , 90° e 135° , conforme mostrado na figura 2.9.

Algumas características a serem calculadas a partir das matrizes GLRL são propostas por Galloway (1975), Chu et al. (1990), Tang (1998), e, apesar de Loh et al. (1988) apresentarem seis medidas, apenas três delas diferem significativamente daquelas descritas nos outros

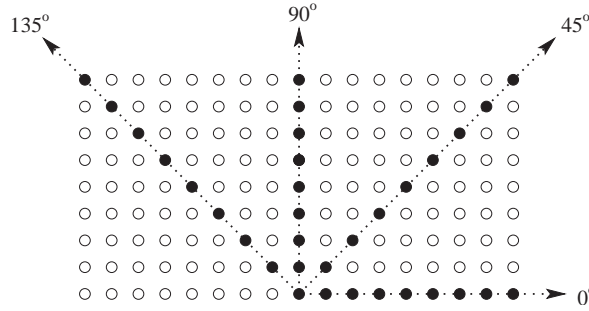


Figura 2.9: Ângulos utilizados para construção das matrizes de comprimento de corridas de cinza.

trabalhos. Na formulação dessas medidas, H_g e N_r representam, respectivamente, o valor do nível de cinza máximo e o tamanho da corrida mais longa.

As medidas denominadas ênfase em corridas curtas (SRE, *short runs emphasis*) e ênfase em corridas longas (LRE, *long runs emphasis*) têm como principal objetivo ressaltar, respectivamente, as corridas curtas e longas para um específico nível de cinza. Portanto, valores maiores de LRE são obtidos quando as corridas encontradas na textura forem maiores, o caso contrário ocorre para SRE. As equações 2.38 e 2.39 apresentam tais medidas.

$$SRE = \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} \frac{P(i, j|\theta)}{j^2} / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.38)$$

$$LRE = \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} j^2 P(i, j|\theta) / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.39)$$

Para descrever como se apresentam as distribuições dos tons de cinza e dos tamanhos de corridas utiliza-se as medidas denominadas não uniformidade de tom de cinza (GLN, *gray level non-uniformity*) e não uniformidade de tamanho de corrida (RLN, *run length non-uniformity*), respectivamente. A característica GLN, mostrada na equação 2.40, apresenta valores baixos quando as corridas são distribuídas de forma homogênea nos tons de cinza, enquanto que a uniformidade na distribuição do tamanho das corridas é responsável por valores baixos para medida RLN, mostrada na equação 2.41.

$$GLN = \sum_{i=0}^{H_g} \left(\sum_{j=1}^{N_r} P(i, j|\theta) \right)^2 / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.40)$$

$$RLN = \sum_{i=1}^{N_r} \left(\sum_{j=0}^{H_g} P(i, j|\theta) \right)^2 / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.41)$$

A última medida proposta por Galloway é denominada percentagem de corrida (RP, *run*

percentage), mostrada na equação 2.42, onde n denota o número de pixels contidos na imagem. Essa medida apresenta valores altos, próximos de um, quando ocorrem apenas corridas curtas.

$$RP = \frac{1}{n} \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.42)$$

Chu et al. (1990) apresentam duas novas características a serem calculadas a partir das matrizes GLRL, denominadas LGRE (*low gray level runs emphasis*) e HGRE (*high gray levels runs emphasis*), mostradas nas equações 2.43 e 2.44, respectivamente. Tais medidas são utilizadas para distinguir texturas que apresentam similaridade em relação aos valores das características SRE e LRE mas diferem na distribuição das corridas dentre os níveis de cinza.

$$LGRE = \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} \frac{P(i, j|\theta)}{i^2} / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.43)$$

$$HGRE = \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} i^2 P(i, j|\theta) / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.44)$$

Visando analisar texturas em cenas tridimensionais, Loh et al. (1988) propõem seis características de texturas a serem obtidas a partir da matriz GLRL, entretanto, apenas três diferem daquelas já apresentadas. Para que essas medidas sejam independentes da escala e orientação da superfície, inicialmente é calculado o fator de normalização NF mostrado na equação 2.45, onde K é uma constante dependente do experimento a ser efetuado e ARL denota o tamanho médio das corridas de cinza.

$$NF = \frac{K}{ARL} \quad (2.45)$$

onde
$$ARL = \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} j P(i, j|\theta) / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta)$$

O segundo momento com respeito ao tamanho da corrida (equação 2.46) e o segundo momento em relação ao nível de cinza (equação 2.47) estão entre as características propostas por Loh et al. (1988). Assim como a medida LRE, definida na equação 2.39, a primeira delas ressalta a presença de corridas longas.

$$F1 = \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} (j \text{ NF})^2 P(i, j|\theta) / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.46)$$

$$F2 = \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} i P(i, j|\theta) / \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} P(i, j|\theta) \quad (2.47)$$

A última medida que apresenta diferença significativa daquelas descritas anteriormente é denominada soma da variância, mostrada na equação 2.48. Essa característica apresenta valores baixos quando ocorrem corridas com o mesmo tamanho para os mesmos tons de cinza.

$$F3 = \sum_{i=0}^{H_g} \sum_{j=1}^{N_r} (j \text{ NF} - U_i)^2 P(i, j|\theta) \quad (2.48)$$

$$\text{onde } U_i = \sum_{j=1}^{N_r} (j \text{ NF}) P(i, j|\theta) / \sum_{j=1}^{N_r} P(i, j|\theta)$$

Utilizando algumas variações na matriz original, Tang (1998) propõe um método de análise de texturas que utiliza os próprios elementos da matriz como características. Devido à existência de um número pequeno de corridas compostas por grande quantidade de pixels, a região da matriz que representa as corridas longas contém poucos elementos. Para que as corridas curtas não sejam demasiadamente ressaltadas, Tang propõe que a matriz GLRL seja criada a partir da equação 2.49, proporcionando igual ênfase para todas as corridas.

$$P_p(i, j|\theta) = j \cdot P(i, j|\theta) \quad (2.49)$$

Além das alterações efetuadas na matriz original, Tang (1998) propõe a criação dos vetores \mathbf{p}_g , definido na equação 2.50, e \mathbf{p}_r , definido na equação 2.51. Tais vetores permitem que haja redução no custo computacional para o cálculo das medidas apresentadas nas equações 2.38 a 2.44, pois representam a distribuição da soma do número de corridas para o i -ésimo tom de cinza e a soma da distribuição de número de corridas com tamanho j , respectivamente.

$$p_g(i) = \sum_{j=1}^{N_r} P(i, j|\theta) \quad i = 0, 1, \dots, H_g \quad (2.50)$$

$$p_r(j) = \sum_{i=0}^{H_g} P(i, j|\theta) \quad j = 1, 2, \dots, N_r \quad (2.51)$$

Dado que as corridas tendem a ser curtas, ou seja, apresentar elementos nas primeiras

colunas, Tang define o *vetor de corrida tamanho um*, mostrado na equação 2.52. Esse vetor aproxima as informações contidas em toda a matriz, proporcionando redução no custo computacional e na dimensionalidade do espaço de características.

$$p_o(i) = p(i, 1) \quad i = 0, 1, \dots, H_g \quad (2.52)$$

Tendo definido os vetores e a matriz, Tang apresenta uma abordagem para extração de características distinta daquelas anteriormente descritas, neste caso, ao invés de extrair medidas, os próprios elementos da matriz e dos vetores são utilizados como características, entretanto, selecionando apenas os mais representativos. Para isso, primeiro executa-se a transformada KLT (Theodoridis e Koutroumbas 2003), visando à redução da dimensionalidade dos dados, e em seguida utiliza-se a *distância de Bhattacharyya* para seleção de características.

Função de Autocorrelação

Uma maneira de diferenciar texturas ásperas de texturas finas baseia-se na detecção da frequência espacial, determinada pela ocorrência de variações de nível de cinza em uma região específica. Texturas finas são compostas por primitivas com tamanho pequeno e apresentam frequência espacial alta devido ao grande número de variações de níveis de cinza, enquanto as texturas ásperas, compostas por primitivas maiores, são caracterizadas por possuírem frequência espacial baixa.

A função de autocorrelação descreve as interações espaciais entre as primitivas. Neste caso, os níveis de cinza dos pixels são considerados como as primitivas que compõem a textura e as interações entre essas primitivas são caracterizadas pelo coeficiente de correlação, que mede a dependência linear entre um pixel e outro separados por uma dada distância.

A equação 2.53 mostra como o coeficiente de autocorrelação é obtido para uma textura composta por $M \times N$ pixels, para cada valor de (p, q) a textura executa um deslocamento de p pixels no eixo x e q pixels no eixo y , como ilustra a figura 2.10. Dessa maneira, os coeficientes $\rho_{ff}(p, q)$ podem ser utilizados como características de texturas. Por exemplo, a variação de (p, q) entre $(0, 0)$ e $(9, 9)$ resulta em um conjunto composto por 100 coeficientes.

$$\rho_{ff}(p, q) = \frac{\sum_{i=0}^{M-p-1} \sum_{j=0}^{N-q-1} f(i, j)f(i + p, j + q)}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f^2(i, j)} \quad (2.53)$$

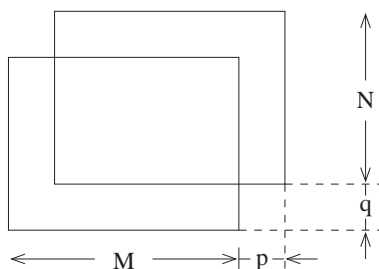


Figura 2.10: Ilustração da função de autocorrelação para uma textura composta por $M \times N$ pixels.

Haralick (1979) descreve a função de autocorrelação como uma característica que mede o tamanho das primitivas tonais que, em primitivas com tamanho grande, o valor do coeficiente de autocorrelação decresce lentamente, enquanto para primitivas pequenas, este coeficiente diminui rapidamente. A figura 2.11 apresenta gráficos da função de autocorrelação para deslocamentos efetuados apenas em p .

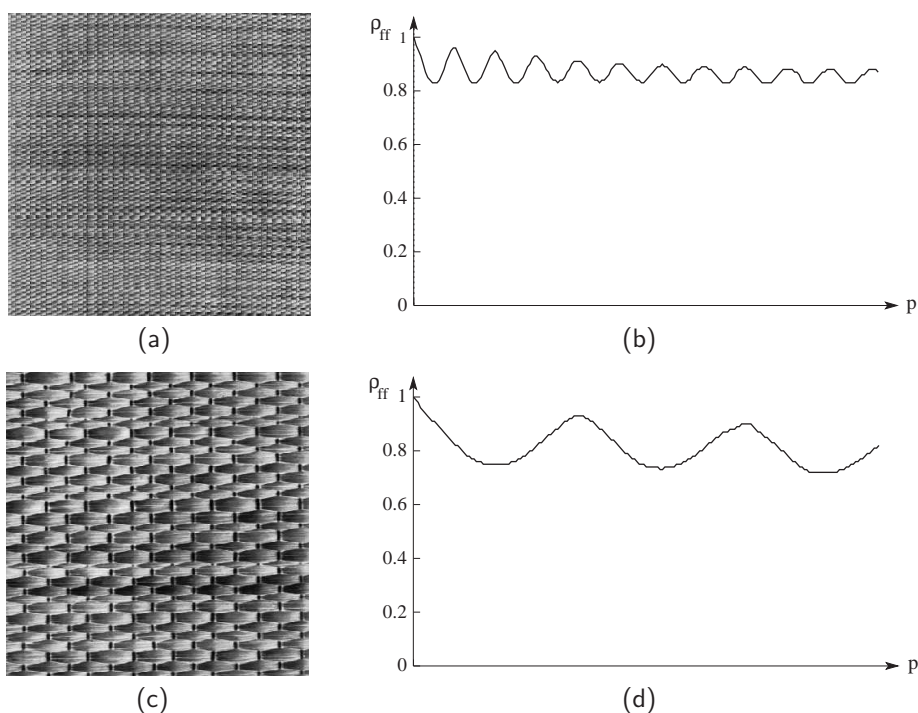


Figura 2.11: Função de autocorrelação para $\rho_{ff}(p, 0)$.

2.2.2 Abordagem Baseada em Processamento de Sinais

Os métodos de análise de texturas baseados em processamento de sinais possuem a característica de extraírem descritores a partir da representação obtida após a execução de transformações na imagem de entrada. A transformada de Fourier, por exemplo, que efetua uma transformação linear nos dados de entrada, possui informações sobre o espectro resultante, essas utilizadas como características de texturas.

Decomposição de Valor Singular

A técnica de decomposição de valor singular (SVD, *singular value decomposition*) apresenta-se como uma importante ferramenta em áreas como restauração de imagens e compressão de dados, nesta última, devido a sua característica de resultar em excelente compactação da energia, propriedade também explorada na análise de texturas.

O método de SVD consiste em decompor uma matriz, cujos elementos podem ser compostos pela intensidade dos pixels pertencentes a uma dada textura, em uma multiplicação de matrizes. Os valores singulares, obtidos com essa decomposição, e sua distribuição provêm informações úteis sobre a textura. Por exemplo, em texturas que apresentam maior aleatoriedade, a energia apresenta-se distribuída entre os valores singulares, enquanto naquelas que se apresentam de forma mais uniforme, a energia fica concentrada apenas em alguns valores singulares.

A partir de resultados obtidos pela álgebra linear, uma matriz \mathbf{A} , $M \times N$ e com $M > N$, pode ser decomposta como mostrado na equação 2.54. Nessa decomposição, denominada decomposição de valor singular, as matrizes \mathbf{U} e \mathbf{V} apresentam, respectivamente, $M \times N$ e $N \times N$ elementos e a diagonal principal da matriz $\mathbf{\Lambda}$, $N \times N$, contém os valores singulares, esses utilizados na determinação das características de texturas.

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T \quad (2.54)$$

Como características a serem extraídas a partir dos valores singulares, podem ser utilizadas as medidas de média e desvio padrão mostradas nas equações 2.55 e 2.56, respectivamente. Nessas equações, $\lambda_k(x, y)$ denota o elemento contido na posição (k, k) da matriz $\mathbf{\Lambda}$, obtida a partir da decomposição singular efetuada sobre uma janela centrada na posição (x, y) da textura.

$$\mu_k = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \lambda_k(i, j) \quad (2.55)$$

$$\sigma_k = \sqrt{\frac{1}{MN-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (\lambda_k(i, j) - \mu_k)^2} \quad (2.56)$$

Após a obtenção das decomposições singulares para cada uma das janelas, um vetor de características composto pelas medidas resultantes pode ser criado para descrever a textura. Por exemplo, se cada janela amostrada contiver 10×8 pixels, oito valores singulares serão

obtidos, desta maneira, a primeira metade do vetor de características pode ser composta pela média e a segunda pelo desvio padrão dos valores singulares.

Espectro de Fourier

O espectro resultante da transformada bidimensional de Fourier, quando efetuado o deslocamento do plano de freqüências da origem, apresenta grande concentração de energia no centro para imagens que possuem componentes de baixa freqüência, enquanto que essa energia fica mais espalhada, localizando-se inclusive em regiões distantes da origem, em imagens de alta freqüência (Gonzalez e Woods 2000). A equação 2.57 representa a transformada de Fourier para uma imagem $n \times n$, onde $i = \sqrt{-1}$.

$$\mathcal{F}(u, v) = \frac{1}{n^2} \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} f(k, l) \exp(-2\pi i(ku + lv)/n) \quad (2.57)$$

Aplicando esses conceitos em texturas, tem-se que o espectro de Fourier para texturas ásperas apresenta concentração de energia no centro do plano, devido à homogeneidade nelas presente. No entanto, para texturas finas, a energia do espectro encontra-se espalhada pelo plano de freqüências. Desse modo, o espectro de Fourier pode ser utilizado como uma característica que avalia a aspereza de texturas.

Expressando-se o espectro em coordenadas polares, levando a uma função $S(r, \theta)$, sendo que S é uma função de espectro, r e θ são variáveis nesse sistema de coordenadas. Para cada direção θ , $S(r, \theta)$ pode ser considerado uma função $S_\theta(r)$. De maneira similar, para cada r , $S_r(\theta)$ é uma função unidimensional. A análise de $S_\theta(r)$ para um valor fixo de θ fornece o comportamento do espectro (como a presença de picos) ao longo de uma direção radial a partir da origem, enquanto a análise de $S_r(\theta)$ para um valor fixo de r leva ao comportamento ao longo de uma circunferência centrada na origem.

Uma descrição global é obtida através das funções mostradas nas equações 2.58 e 2.59, onde R denota o raio de uma circunferência centrada na origem, para um espectro de $n \times n$. O valor de R é tipicamente escolhido como $n/2$.

$$S(r) = \sum_{\theta=0}^{\pi} S_\theta(r) \quad (2.58)$$

$$S(\theta) = \sum_{r=1}^R S_r(\theta) \quad (2.59)$$

Os resultados das equações 2.58 e 2.59 constituem um par de valores $[S(r), S(\theta)]$ para

cada par de coordenadas (r, θ) . Variando-se essas coordenadas, pode-se gerar funções unidimensionais $S(r)$ e $S(\theta)$ que constituem descrições de energia espectral da textura para uma imagem ou região em questão.

Descritores tipicamente usados para esse propósito são a posição do valor mais alto, a média e a variância da amplitude e as variações de eixo, e a distância entre a média e o maior valor da função. A figura 2.12 ilustra o uso da equação 2.59 para a descrição global de textura, dado que cada imagem apresenta primitivas com tamanho e orientações distintas, percebe-se diferenças claras entre as funções $S(\theta)$ resultantes.

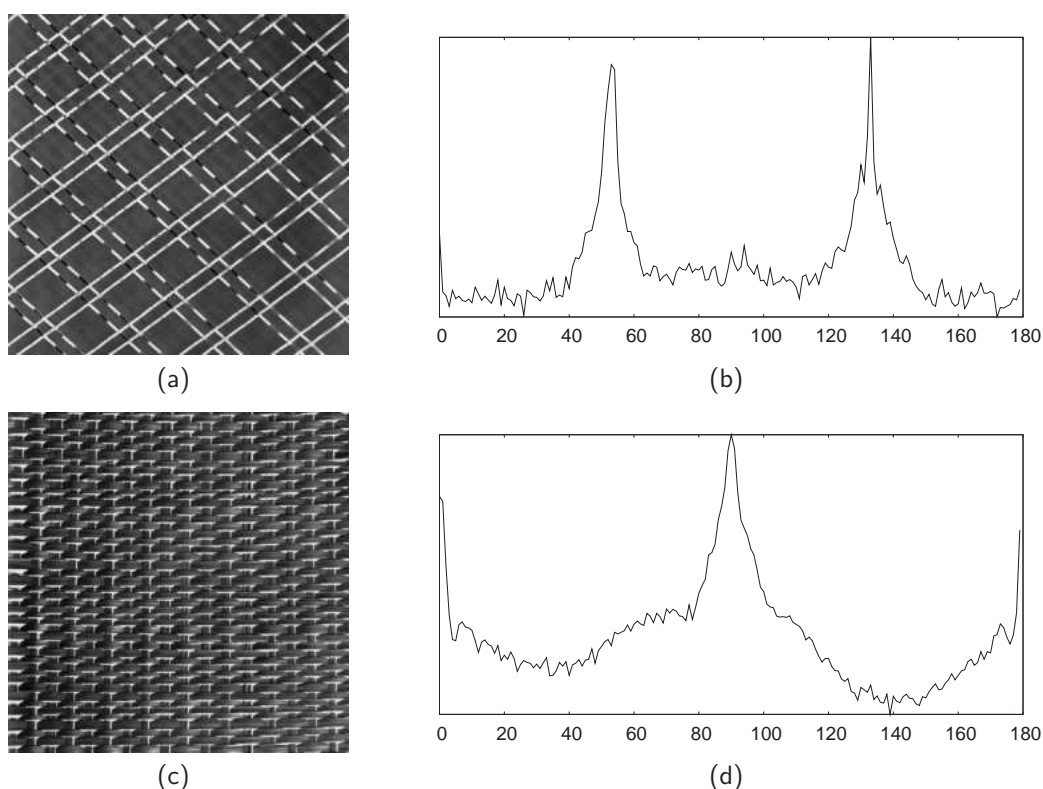


Figura 2.12: Funções $S(\theta)$ determinadas a partir do espectro de Fourier das texturas mostradas em (a) e (c). O eixo das abscissas representa o parâmetro θ .

Transformada Wavelet

Técnicas que aplicam multi-resolução em imagens, tais como *wavelets*, objetivam alterar a representação de modo que tanto informações sobre frequência quanto informações espaciais estejam presentes. O uso da transformada wavelet foi proposta inicialmente para análise de texturas no trabalho de Mallat (1989).

Essa transformada decompõe um sinal por meio de uma série de funções elementares, chamadas *wavelets* e *escala*, criadas por escalamentos e translações de uma função de base, deno-

minada wavelet mãe, mostrada na equação 2.60, onde s guia o escalamento e u a translação.

$$\Psi_{s,u}(x) = \frac{1}{\sqrt{s}} \Psi\left(\frac{x-u}{s}\right) \quad (2.60)$$

A decomposição de uma função pode ser obtida por meio de convoluções entre as funções elementares wavelets e escala com o sinal de entrada, como mostra equação 2.61, onde $f(x)$ contém a função do sinal a ser decomposto.

$$Wf(s, u) = \int_{-\infty}^{\infty} f(x) \Psi_{s,u}^*(x) dx \quad (2.61)$$

Quando a transformada é aplicada em imagens bidimensionais, as funções wavelets são aplicadas como filtros passa-alta, enquanto as funções de escala atuam como filtros passa-baixa. Este processo é executado recursivamente sempre para a sub-imagem que apresenta menor frequência. Ao final da aplicação, a imagem original apresenta-se decomposta em uma série de imagens com escalas e frequências distintas. A figura 2.13 ilustra a decomposição efetuada pela transformada wavelet.

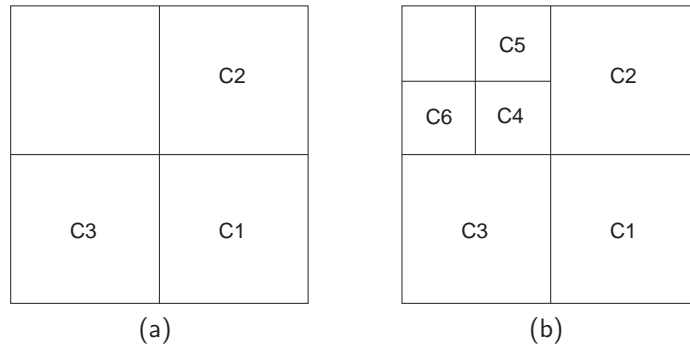


Figura 2.13: Decomposição resultante da aplicação da transformada wavelet. (a) decomposição em um nível; (b) decomposição em dois níveis.

Desde que as imagens de baixa frequência derivadas da decomposição perdem informações importantes sobre a textura (Ruiz et al. 2004), apenas as sub-imagens resultantes de aplicações dos filtros passa-alta são utilizadas para extração de características de texturas. Dentre essas características, encontra-se a energia de uma sub-imagem (Chang e Kuo 1993, de Wouwer et al. 1999), mostrada na equação 2.62, onde $x(l, m)$ representa os coeficientes resultantes da transformada, E_{C_i} denota a energia da i -ésima sub-imagem, conforme mostrado na figura 2.13(b), e $n_s \times n_s$ apresenta as dimensões da sub-imagem em questão.

$$E_{C_i} = \frac{1}{n_s^2} \sum_{l=0}^{n_s-1} \sum_{m=0}^{n_s-1} x(l, m)^2 \quad i = 1, \dots, 6 \quad (2.62)$$

Outras características de texturas a serem extraídas a partir da decomposição obtida pela transformada wavelet são estatísticas de segunda ordem calculadas por meio de matrizes de co-ocorrência sobre as sub-imagens que apresentam alta frequência e as medidas obtidas a partir de histogramas, neste caso resultando em estatísticas de primeira ordem (de Wouwer et al. 1999).

2.2.3 Abordagem Geométrica

Na abordagem geométrica, uma textura é definida como sendo composta por primitivas, também chamadas *textels* (*texture elements*). Após a identificação das primitivas que compõem a textura, duas classes de métodos são utilizadas para extração de características. A primeira utiliza medidas extraídas das primitivas para descrever a textura, enquanto a segunda extrai regras para descrever a disposição espacial dessas primitivas.

A vantagem de se utilizar a segunda classe de métodos, denominados estruturais, está no fato que eles provêm uma boa descrição simbólica da textura, característica bastante útil para síntese de texturas. Por outro lado, na análise de texturas os métodos estruturais não apresentam boa adaptação, pois apenas atuam de maneira satisfatória em texturas regulares, as quais praticamente não ocorrem em imagens naturais. Dado que o foco deste trabalho é na análise de texturas, o restante desta seção é destinada aos métodos que extraem medidas a partir das primitivas de texturas.

Unidade de Textura

He e Wang (1990) propõem o conceito de *unidade de textura*, baseado na idéia de que uma imagem texturizada pode ser considerada como um conjunto de pequenas unidades essenciais, denominada unidade de textura, as quais caracterizam a informação local de um dado pixel em relação aos seus vizinhos. Medidas extraídas a partir de todas as unidades presentes na imagem revelam o aspecto global da textura.

Seja uma vizinhança de 3×3 pixels, composta pelos elementos $V = \{v_0, v_1, \dots, v_8\}$, onde v_0 representa o tom de cinza do pixel central e os outros v_i os tons de cinza de seus vizinhos. Define-se unidade de textura (TU, *texture unit*) pelo conjunto $TU = \{e_1, e_2, \dots, e_8\}$, onde cada e_i é determinado através da equação 2.63.

$$e_i = \begin{cases} 0, & \text{se } v_i < v_0 \\ 1, & \text{se } v_i = v_0 \\ 2, & \text{se } v_i > v_0 \end{cases} \quad (i = 1, 2, \dots, 8) \quad (2.63)$$

Baseado no fato de existirem 6561 ($= 3^8$) configurações possíveis para cada unidade de textura, cria-se uma *assinatura* conforme define a equação 2.64. A distribuição de frequências das unidades de uma textura é denominada *espectro de textura*, onde o eixo das abscissas indica o N_{TU} , denominado número da unidade de textura, o eixo das ordenadas representa seu número de ocorrências.

$$N_{TU} = \sum_{i=1}^8 3^{i-1} e_i \quad (2.64)$$

As figuras 2.14(b) e (d) apresentam o espectro das texturas mostradas nas figuras 2.14(a) e (c), respectivamente. A partir dessa distribuição de frequências são extraídas as características de texturas, que acrescentadas em um vetor de características são utilizadas como discriminantes em métodos de segmentação ou classificação de regiões.

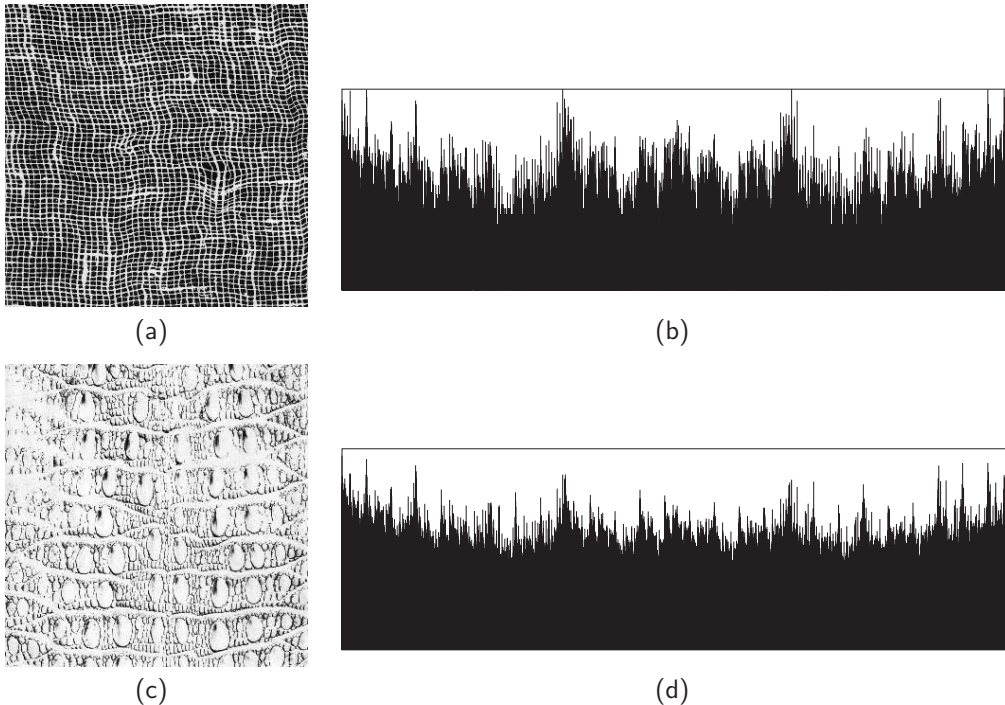


Figura 2.14: Representação do espectro de textura. (a), (c) imagens originais; (b), (d) respectivos espectros de textura representados em escala logarítmica.

A equação 2.64 define como deve ser calculada a unidade de textura para cada conjunto de pixels selecionado, no entanto, exceto para o v_0 , não especifica quais pixels correspondem a

cada v_i . Se considerar apenas ordenações no sentido horário, um elemento v_i , com i fixo, pode assumir oito posições distintas, por exemplo, de a até h do diagrama mostrado na figura 2.15. Portanto, as 6561 unidades de textura podem ser rotuladas, por meio da equação 2.64, de oito maneiras distintas.

a	b	c
h		d
g	f	e

Figura 2.15: Vizinhança utilizada para determinação da unidade de textura.

Utilizando o conceito de unidade de textura, He e Wang (1991) apresentam cinco medidas para características texturais, a simetria preto-branco, simetria geométrica, grau de direção, atributos de orientação e simetria central. A medida de simetria preto-branco (BWS, *black-white symmetry*) retorna valores normalizados entre 0 e 100, onde altos valores de BWS mostram que a inversão dos valores dos pixels da imagem original não altera o seu espectro de textura. O cálculo da medida BWS dá-se pela equação 2.65, em que $S(i)$ contém a frequência com que aparecem unidades com o valor de *assinatura* i no espectro de textura.

$$BWS = \left[1 - \frac{\sum_{i=0}^{3279} |S(i) - S(3281 + i)|}{\sum_{i=0}^{6560} S(i)} \right] \times 100 \quad (2.65)$$

A simetria geométrica (GS, *geometric symmetry*) apresenta informações sobre a regularidade da forma da textura. Ela mede a simetria entre as frequências das regiões a e e , b e f , c e g , d e h , definidas pela figura 2.15. Altos valores de GS mostram que a rotação de 180 graus na imagem não altera o espectro de textura. O GS é calculado pela equação 2.66, $S_j(i)$ contém a frequência com que ocorrem unidades com o valor de *assinatura* i sob a ordenação j , onde j representa uma dentre as oito possíveis ordenações utilizadas para o cálculo da unidade de textura.

$$GS = 100 \left[1 - \frac{\frac{1}{4} \sum_{j=1}^4 \sum_{i=0}^{6560} |S_j(i) - S_{j+4}(i)|}{2 \sum_{i=0}^{6560} S_j(i)} \right] \quad (2.66)$$

Outra medida definida sobre o espectro de textura é o grau de direção (DD, *degree of direction*), responsável por medir o grau de linearidade das estruturas presentes em uma textura.

Valores altos de DD indicam que o espectro de textura é sensível com relação ao padrão de orientação da imagem. A equação 2.67 mostra como a medida DD é calculada, onde $S_m(i)$ e $S_n(i)$ são definidos da mesma maneira que $S_j(i)$ na equação 2.66.

$$DD = 100 \left[1 - \frac{1}{6} \sum_{m=1}^3 \sum_{n=m+1}^4 \frac{\sum_{i=0}^{6560} |S_m(i) - S_n(i)|}{2 \sum_{i=0}^{6560} S_m(i)} \right] \quad (2.67)$$

He e Wang, considerando que as texturas apresentam micro-estruturas direcionadas, definem medidas para descrever as propriedades de orientação, na horizontal (MHS, *micro-horizontal structure*), vertical (MVS, *micro-vertical structure*) e nas duas diagonais, MDS1 (*micro-diagonal structure 1*) e MDS2 (*micro-diagonal structure 2*). As equações 2.68 a 2.71 apresentam as medidas criadas, onde $P(k, l, m)$ denota o número de elementos que possuem valores idênticos para e_k , e_l e e_m na i -ésima unidade de textura.

$$MHS = \sum_{i=0}^{6560} S(i) P(1, 2, 3) P(6, 7, 8) \quad (2.68)$$

$$MVS = \sum_{i=0}^{6560} S(i) P(1, 4, 6) P(3, 5, 8) \quad (2.69)$$

$$MDS1 = \sum_{i=0}^{6560} S(i) P(4, 1, 2) P(7, 8, 5) \quad (2.70)$$

$$MDS2 = \sum_{i=0}^{6560} S(i) P(2, 3, 5) P(4, 6, 7) \quad (2.71)$$

A última característica de textura definida por He e Wang, denominada simetria central (SC), baseia-se nas relações entre as orientações (a, h) , (b, g) , (c, f) e (d, e) , definidas no diagrama mostrado na figura 2.15. A equação 2.72 mostra como é efetuado este cálculo, $K(i)$ representa o número de pares que apresentam os mesmos valores nos elementos (e_a, e_h) , (e_b, e_g) , (e_c, e_f) e (e_d, e_e) , onde e_i é determinado pela equação 2.63.

$$SC = \sum_{i=0}^{6560} S(i) [K(i)]^2 \quad (2.72)$$

Como alternativa ao N_{TU} , definido na equação 2.64, Ojala et al. (1996) apresentam a versão binária da unidade de textura, denominada padrões locais binários (LBP, *local binary*

patterns), seguindo a definição mostrada na equação 2.73.

$$e_i = \begin{cases} 0, & \text{se } v_i < v_0 \\ 1, & \text{se } v_i \geq v_0 \end{cases} \quad (2.73)$$

Considerando a alteração na definição dos elementos e_i , a equação 2.74 passa a ser utilizada para o cálculo da LBP. Dessa maneira, o valor máximo que o LBP pode assumir é 255, efetuando uma redução significativa no número de entradas do espectro de textura. Combinando os valores de LBP com medidas de contraste e covariância, Pietikäinen et al. (2000) obtêm a distribuição conjunta das características para cada textura.

$$\text{LBP} = \sum_{i=1}^8 2^{i-1} e_i \quad (2.74)$$

Outra alteração efetuada sobre a unidade de textura é proposta por Al-Janobi (2001) que, ao invés de modificar a definição de e_i , como propuseram Ojala et al., altera a estrutura da unidade de textura. Essa modificação, denominada matriz de textura diagonal-e-cruz (CDTM, *cross-diagonal texture matrix*), particiona a unidade de textura em dois conjuntos, o primeiro, denominado CTU, composto pelos vizinhos na vertical e horizontal em relação ao pixel central e o segundo composto pelos elementos das diagonais, denominado DTU, como mostrado na figura 2.16.

e_1	e_2	e_3
e_8		e_4
e_7	e_6	e_5

(a)

e_{d_1}		e_{d_2}
e_{d_4}		e_{d_3}

(b)

	e_{c_1}	
e_{c_4}		e_{c_2}
	e_{c_3}	

(c)

Figura 2.16: Decomposição da unidade de textura. (a) estrutura original da unidade de textura; (a) matriz diagonal; (b) matriz cruz.

Com a decomposição efetuada na unidade de textura, Al-Janobi obtêm a distribuição

conjunta proporcionada pelos valores obtidos nas equações 2.75 e 2.76, a partir da CTU e DTU, respectivamente. Finalmente, as medidas estatísticas definidas por Haralick et al. (1973) são utilizadas para extrair informações a partir da distribuição conjunta obtida.

$$N_{CTU} = \sum_{i=1}^4 3^{i-1} e_{Ci} \quad (2.75)$$

$$N_{DTU} = \sum_{i=1}^4 3^{i-1} e_{Di} \quad (2.76)$$

Codificação de Características de Texturas

Horng et al. (2002) propõem o método denominado codificação de características de texturas (TFCM, *texture feature coding method*), utilizando sete medidas para descrição das características. Altera-se o conceito de unidade de textura definido por He e Wang (1990) (seção sobre *unidade de textura*), considerando apenas a sua conectividade, definindo a conectividade-4 de primeira ordem e a conectividade-4 de segunda ordem, como mostra a figura 2.17.

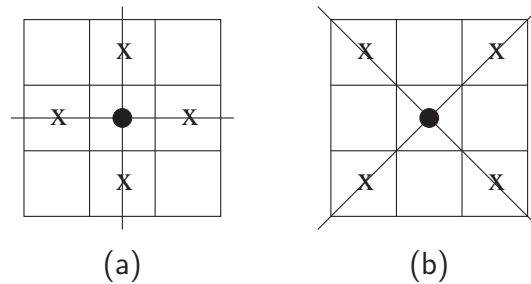


Figura 2.17: Conectividade sobre a unidade de textura. (a) conectividade-4 de primeira ordem; (b) conectividade-4 de segunda ordem.

O TFCM produz pares de inteiros (α, β) , onde α e β representam a variação do nível de cinza da conectividade de primeira e de segunda ordem, respectivamente. A determinação das variações dos tons de cinza é efetuada pela varredura de três píxels consecutivos, conforme as ordens de conectividade definidas.

Considerando (a, b, c) uma ordenação para varredura e (g_a, g_b, g_c) os tons de cinza desses píxels, se considerar duas variações sucessivas entre os pares (g_a, g_b) e (g_b, g_c) , há quatro tipos de variações, conforme definido pela equação 2.77, onde Δ denota a tolerância da variação.

- (i) se $(|g_a - g_b| \leq \Delta)$ e $(|g_b - g_c| \leq \Delta)$
(ii) se $[(|g_a - g_b| \leq \Delta) \text{ e } (|g_b - g_c| \geq \Delta)]$ ou $[(|g_a - g_b| \geq \Delta) \text{ e } (|g_b - g_c| \leq \Delta)]$
(iii) se $[(g_a - g_b > \Delta) \text{ e } (g_b - g_c > \Delta)]$ ou $[(g_b - g_a > \Delta) \text{ e } (g_c - g_b > \Delta)]$
(iv) se $[(g_a - g_b > \Delta) \text{ e } (g_c - g_b > \Delta)]$ ou $[(g_b - g_a > \Delta) \text{ e } (g_b - g_c > \Delta)]$
- (2.77)

Conforme a equação 2.77, a variação do tipo (i) descreve o caso em que a, b e c apresentam valores próximos, dentro da tolerância; no tipo (ii) ocorre a variação do tom de cinza apenas em um par de pixels. O tipo (iii) denota o caso onde acontece variação crescente ou decrescente em relação à tripla de pixels (a, b, c). No tipo (iv), o pixel central apresenta valor maior ou menor que seus vizinhos.

Como cada ordem de conectividade-4 apresenta duas varreduras, criam-se valores para determinação de α e β para cada pixel (x, y) da imagem, conforme mostrado na figura 2.18. Define-se o número da característica de textura (NCT) como $NCT(x, y) = \alpha(x, y)\beta(x, y)$, esta multiplicação não produz todos os números entre 0 e 100, portanto, compacta-se o resultado mapeando-os para valores entre 0 e 41.

	(i)	(ii)	(iii)	(iv)
(i)	1	2	3	4
(ii)	2	5	6	7
(iii)	3	6	8	9
(iv)	4	7	9	10

Figura 2.18: Valores utilizados para determinação de α e β no método TFCM para análise de texturas. As linhas indicam o tipo da primeira varredura, as colunas indicam o tipo da segunda varredura.

Após a determinação do NCT para todos os pixels da imagem, calcula-se um histograma e a matriz de co-ocorrência (seção 2.2.1), ambos utilizados para a determinação das medidas que descrevem as texturas analisadas através do método TFCM. O histograma é calculado pela equação 2.78, $N(i)$ denota o número de ocorrências do i-ésimo NCT e n representa o número de pixels presentes na imagem. A matriz de co-ocorrência é calculada como mostra a equação 2.79, $N_{d,\theta}(i, j)$ denota o número de transições de NCT i para j com distância d e ângulo θ , e N_t representa o número total de transições.

$$p_{TFCM}(i) = \frac{N(i)}{n}, \quad n \in \{0, 1, \dots, 41\} \quad (2.78)$$

$$p_{TFCM}(i, j|d, \theta) = \frac{N_{d,\theta}(i, j)}{N_t}, \quad i, j \in \{0, 1, \dots, 41\} \quad (2.79)$$

A descrição das características das texturas analisadas com o método TFCM dá-se mediante sete medidas baseadas no histograma e na matriz de co-ocorrência calculados sobre os valores de NCT. São elas, aspereza (As), homogeneidade (Hom), convergência média (CM), variância (Va), entropia de código (CE), similaridade de código (CS) e similaridade de resolução (RS).

As medidas de aspereza, homogeneidade, convergência média e variância são baseadas no histograma. Valores altos de aspereza (equação 2.80) indicam alterações drásticas nos valores dos pixels¹, exatamente o contrário do que indicam valores altos para homogeneidade (equação 2.81). A medida de convergência média (equação 2.82) descreve quão próximo a textura está da média, onde μ e σ denotam a média e o desvio padrão dos NCT, respectivamente. A variância (equação 2.83) mede o desvio dos valores de NCT da média.

$$As = \sum_x \sum_y p_{TFCM}(41) \quad (2.80)$$

$$Hom = \sum_x \sum_y p_{TFCM}(0) \quad (2.81)$$

$$CM = \sum_{n=0}^{41} \frac{|n p_{TFCM}(n) - \mu|}{\sigma} \quad (2.82)$$

$$Var = \sum_{n=0}^{41} (n - \mu)^2 p_{TFCM}(n) \quad (2.83)$$

A partir da matriz de co-ocorrência são determinadas as medidas de entropia de código e similaridade de código. A entropia de código (equação 2.84) mede a entropia da matriz de co-ocorrência e a similaridade de código (equação 2.85) descreve a densidade dos mesmos NCT para vizinhança 8-conectada.

$$CE = - \sum_{i=0}^{41} \sum_{j=0}^{41} p_{TFCM}(i, j|d, \theta) \log p_{TFCM}(i, j|d, \theta) \quad (2.84)$$

$$CS = \sum_{i=0}^{41} p_{TFCM}^2(i, i|d, \theta) \quad (2.85)$$

A Similaridade de resolução (equação 2.86) é a última medida definida por Horng et al. (2002). Ela provê informações sobre a probabilidade conjunta $p(i, j; x, y)$ do pixel com coordenadas (x, y) que apresenta NCT i quando $\Delta = 0$ e NCT j quando Δ assume valor utilizado

¹Horng et al. (2002) utilizam o termo aspereza, entretanto, para apresentar maior clareza deveriam denominá-la fineza, pois valores altos da medida são caracterizados por alterações drásticas nos níveis de cinza dos pixels.

correntemente. Valores altos de RS indicam que a textura apresenta maior aspereza.

$$RS = \sum_x \sum_y \frac{p(i, j; x, y)}{1 + (i - j)^2} \quad (2.86)$$

2.2.4 Abordagem Baseada em Modelos Paramétricos

Nesta abordagem, uma textura é considerada como uma amostra extraída de um processo estocástico definido por um conjunto de parâmetros que, servindo como modelo para textura, resumem suas características. Com a utilização desses parâmetros é possível tanto efetuar a análise quanto a síntese da textura, essa última por meio de um processo de amostragem.

Esta seção trata dos modelos paramétricos baseados em campos aleatórios de Markov (definido na seção 2.1), descrevendo alguns modelos cujos parâmetros são utilizados para análise de texturas. Chen e Huang (1993) citam que, dentre as principais vantagens de se utilizar MRF, está o fato de que seus parâmetros capturam a direcionalidade e a aspereza da textura, além de não necessitar de uma etapa subsequente de seleção ou nova extração de características.

Campos Aleatórios de Markov

A análise de texturas baseada em campos aleatórios de Markov utiliza os parâmetros obtidos a partir da distribuição de probabilidade das características locais. Por outro lado, se tais parâmetros forem conhecidos, texturas sintéticas podem ser obtidas a partir de métodos de amostragem.

O modelo de características locais proposto por Ising é utilizado por Cross e Jain (1983) para modelagem e síntese de texturas, considerando vizinhanças de primeira, segunda, terceira e quarta ordem. A equação 2.87 mostra a formulação das características locais, onde o parâmetro q (energia da definição 2.9) depende da ordem da vizinhança utilizada, conforme o esquema mostrado pela figura 2.19.

$$P(X_s = x_s | \partial_s) = \frac{\exp(-qx)}{1 + \exp(-q)} \quad (2.87)$$

O modelo de vizinhança de primeira ordem utiliza q com a forma

$$q = a + b_{1,1}(x_t + x_{t'}) + b_{1,2}(x_u + x_{u'}) \quad (2.88)$$

	o_1	m	q_1	
o_2	v	u	z	q_2
l	t	s	t'	l'
q'_1	z'	u'	v'	o'_1
	q'_2	m'	o'_2	

Figura 2.19: Representação da vizinhança do pixel central com valor x_s .

para vizinhança de segunda ordem

$$q = \alpha + b_{1,1}(x_t + x_{t'}) + b_{1,2}(x_u + x_{u'}) + b_{2,1}(x_v + x_{v'}) + b_{2,2}(x_z + x_{z'}) \quad (2.89)$$

para vizinhança de terceira ordem

$$q = \alpha + b_{1,1}(x_t + x_{t'}) + b_{1,2}(x_u + x_{u'}) + b_{2,1}(x_v + x_{v'}) + b_{2,2}(x_z + x_{z'}) \\ + b_{3,1}(x_m + x_{m'}) + b_{3,2}(x_l + x_{l'}) \quad (2.90)$$

e, finalmente, para vizinhança de quarta ordem, q torna-se

$$q = \alpha + b_{1,1}(x_t + x_{t'}) + b_{1,2}(x_u + x_{u'}) + b_{2,1}(x_v + x_{v'}) + b_{2,2}(x_z + x_{z'}) \\ + b_{3,1}(x_m + x_{m'}) + b_{3,2}(x_l + x_{l'}) + b_{4,1}(x_{o_1} + x_{o'_1} + x_{o_2} + x_{o'_2}) \\ + b_{4,2}(x_{q_1} + x_{q'_1} + x_{q_2} + x_{q'_2}) \quad (2.91)$$

Os coeficientes α e $b_{i,j}$ são os parâmetros da distribuição de probabilidade do MRF, utilizados diretamente para descrição da textura. A determinação desses parâmetros deve ser feita com a utilização de estimadores como o método *coding*, proposto por Besag (1974), e a pseudo-máxima verossimilhança.

A aplicação do método da pseudo-máxima verossimilhança para estimação dos parâmetros objetiva a maximização da equação 2.92, onde θ é o vetor com os parâmetros a serem estimados. Esta maximização pode ser feita utilizando-se o método de Newton-Raphson para encontrar múltiplas raízes das derivadas parciais (Ohanian e Dubes 1992). A figura 2.20 mostra a estimação dos parâmetros para duas texturas, utilizando vizinhança de primeira ordem.

$$L(\theta) = \prod_{s \in S} P(X_s = x_s | \partial_s) \quad (2.92)$$

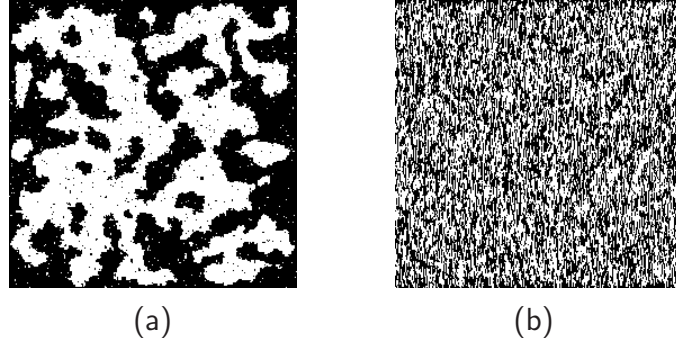


Figura 2.20: Estimação dos parâmetros α , $b_{1,1}$ e $b_{1,2}$ do MRF que modela as texturas. (a) $\alpha = -6.0$, $b_{1,1} = 3.0$ e $b_{1,2} = 3.0$; (b) $\alpha = -2.0$, $b_{1,1} = 0.08$ e $b_{1,2} = 3.0$.

Em seu artigo comparativo, Ohanian e Dubes (1992) utilizam o modelo auto-binomial (Besag 1974) com vizinhança de segunda ordem para o MRF, como mostra a equação 2.93, onde G denota o número de tons de cinza presentes na imagem e $\theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ os parâmetros a serem estimados utilizando-se a pseudo-máxima verossimilhança. Este modelo possibilita a determinação dos parâmetros de texturas com mais de dois tons de cinza, ao contrário do modelo utilizado por Cross e Jain (1983).

$$P(X_s = x_s | \theta_s) = \binom{G-1}{x_s} q^{x_s} (1-q)^{G-1-x_s} \quad (2.93)$$

onde $q = \frac{\exp(-T)}{1 + \exp(-T)}$, $T = \theta_1(x_t + x_{t'}) + \theta_2(x_u + x_{u'}) + \theta_3(x_z + x_{z'}) + \theta_4(x_v + x_{v'})$

2.3 Formulação Bayesiana

A segmentação tem como objetivo particionar uma imagem em regiões homogêneas, atribuindo os pixels pertencentes a regiões com as mesmas características a uma mesma classe. Dessa maneira, considerando a existência de m classes distintas, cada um dos n pixels que compõem a imagem deve pertencer a uma das classes $L = \{1, 2, \dots, m\}$, onde L representa uma especificação da definição 2.1 e a segmentação correta é dada pelo vetor $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$, $x_i^* \in L$.

O conjunto de variáveis observadas é denotado por $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$, onde cada elemento y_i é um vetor de características associado ao i -ésimo pixel. Neste trabalho, a obtenção desses vetores é dada a partir dos métodos de análise de texturas descritos na seção 2.2. As informações sobre dependência local, na qual regiões próximas tendem a assumir valores parecidos, são modeladas por um MRF (descrito na seção 2.1), onde suas variáveis aleatórias são representadas pelo conjunto $\mathbf{X} = \{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}$, $x_i \in L$.

Dado um conjunto de vetores de características \mathbf{y} e as informações sobre a dependência

espacial \mathbf{X} , o problema concentra-se em encontrar a segmentação ótima \mathbf{x}^* , que é uma realização específica de \mathbf{X} . Normalmente, os métodos de segmentação baseiam-se apenas nas informações contidas nos vetores de características (Reed e Dubuf 1993), entretanto, técnicas de segmentação mais acuradas utilizam a formulação Bayesiana (D'Elia et al. 2003, Deng e Clausi 2004, Liapis et al. 2004), combinando os vetores de características com informações sobre a dependência espacial, como mostra a equação 2.94, conhecida como teorema de Bayes, onde $P(\mathbf{y}|\mathbf{x})$ denota a distribuição das características dos pixels, dada a segmentação \mathbf{x} .

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{x})P(\mathbf{y}|\mathbf{x})}{P(\mathbf{y})} \quad (2.94)$$

Essa formulação possibilita a determinação da probabilidade condicional $P(\mathbf{x}|\mathbf{y})$, denominada probabilidade *a posteriori*, através da associação do conhecimento obtido inicialmente com os vetores de características, e as informações sobre a dependência espacial, modelada por um MRF com a medida de probabilidade $P(\mathbf{x})$, denominada probabilidade *a priori*. Diferentemente do uso tradicional da formulação Bayesiana, a dependência espacial, representada por $P(\mathbf{x})$, é atualizada durante o processo de segmentação, pois depende das relações locais, conforme descrito na seção 2.1.

Para que a probabilidade *a posteriori* $P(\mathbf{x}|\mathbf{y})$ seja um MRF, considera-se a independência condicional das variáveis aleatórias que compõem a distribuição de características dos pixels, dada a segmentação \mathbf{x} . Assumindo essa condição, Dubes e Jain (1989) apresentam a equação 2.95 como sendo a distribuição *a posteriori*, onde $H(\mathbf{x}|\mathbf{y})$ denota a função de energia, mostrada na definição 2.9, e Z representa a função de partição.

$$P(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} \exp\{-H(\mathbf{x}|\mathbf{y})\} \quad (2.95)$$

Com a equação 2.95, dados os vetores de características e informações sobre a dependência espacial, o problema recai na estimação ótima da segmentação correta \mathbf{x}^* , a qual é efetuada com a utilização do método MAP (*maximum a posteriori*), que seleciona a realização \mathbf{x} que maximiza a probabilidade *a posteriori* $P(\mathbf{x}|\mathbf{y})$.

Entretanto, o custo computacional para determinar a segmentação ótima é extremamente alto, devido à necessidade de calcular m^n vezes a probabilidade condicional $P(\mathbf{x}|\mathbf{y})$. Dessa maneira, para que a segmentação seja efetuada em tempo computacional aceitável, pode-se utilizar alguns algoritmos que aproximam a obtenção de \mathbf{x}^* pela estimação MAP.

Para obter uma estimação de \mathbf{x}^* , esses algoritmos utilizam técnicas de relaxação, processo que tenta diminuir a quantidade de energia presente em um sistema, o que equivale a maxi-

mizar a probabilidade de se obter uma segmentação correta (Geman e Geman 1984). Dentre os principais métodos estão o resfriamento simulado (SA, *simulated annealing*), proposto por Geman e Geman (1984), a maximização de probabilidades marginais (MPM, *maximizer of posterior marginals*) (Marroquin et al. 1987) e o modos condicionais de iteração, proposto por Besag (1986).

Geman e Geman (1984) propõem um método para maximizar a probabilidade *a posteriori* utilizando resfriamento simulado (Kirkpatrick et al. 1983), que se baseia no algoritmo de Metropolis et al. (1953). O objetivo desse método é maximizar iterativamente a probabilidade $P(\mathbf{x}|\mathbf{y})$ (equivalente a minimizar a função de energia $H(\mathbf{x}|\mathbf{y})$) por meio de trocas locais nas variáveis aleatórias, como mostra o algoritmo 2.1. Este algoritmo possibilita a obtenção de estados que apresentam maior quantidade de energia, evitando a permanência em regiões de mínimos locais.

Estimação MAP utilizando *simulated annealing*

Sejam $\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^k, \dots$ realizações pertencentes ao espaço amostral $\Omega = \{1, 2, \dots, m\}^n$ criadas no instante k , onde $f(T_k)$ é uma função monotonicamente decrescente, $H(\mathbf{x}^k|\mathbf{y})$ denota a função de energia do MRF e o parâmetro T_k a temperatura no instante k .

```

1 // inicialização
2  $T_0 \leftarrow$  temperatura inicial
3  $\mathbf{x}_i^0 \leftarrow$  valoração aleatória //  $\forall x_i \in \mathbf{x}$ 
4
5 // escolha de uma variável aleatória e um valor
6  $v \leftarrow$  escolha aleatória em  $\{1, 2, \dots, m\}$ 
7  $s \leftarrow$  escolha aleatória em  $\{1, 2, \dots, n\}$ 
8
9 // cria a nova realização, no instante  $k+1$ 
10  $\mathbf{z} = (X_1^k = x_1^k, X_2^k = x_2^k, \dots, X_s^k = v, \dots, X_n^k = x_n^k)$ 
11
12 // escolha da realização utilizada no instante  $k+1$ 
13  $\Delta H \leftarrow H(\mathbf{z}|\mathbf{y}) - H(\mathbf{x}^k|\mathbf{y})$ 
14 se  $\Delta H < 0$ 
15      $\mathbf{x}^{k+1} \leftarrow \mathbf{z}$ 
16 senão
17      $\mathbf{x}^{k+1} \leftarrow \mathbf{z}$  com probabilidade  $e^{-\Delta H/T_k}$ 
18
19 repita os passos 6–17 por  $200n$  vezes
20  $T_{k+1} \leftarrow f(T_k)$ 
21 repita os passos 6–19 até  $T_k \approx 0$ 

```

Algoritmo 2.1: Estimação MAP utilizando *simulated annealing*.

A idéia central do método é gerar uma seqüência de realizações e, conforme o parâmetro T

vai atingindo valores menores, a quantidade de energia vai diminuindo e o valor da probabilidade *a posteriori* aumenta. Isso pode ser percebido através da equação 2.95, se o valor de $H(\mathbf{x}|\mathbf{y})$ decresce, o valor de $P(\mathbf{x}|\mathbf{y})$ aumenta. No entanto, a diminuição da temperatura deve ser feita lentamente (Kirkpatrick et al. 1983). Geman e Geman utilizam a equação 2.96 como guia para o decaimento da temperatura, Manjunath et al. (1990) propõem a equação 2.97 e Blake (1989) apresenta a equação 2.98.

$$T_{k+1} = \frac{\ln(1+k)}{\ln(2+k)} T_k \quad (2.96)$$

$$T_{k+1} = \frac{T_0}{1 + \log_2(k)} \quad (2.97)$$

$$T_{k+1} = \frac{T_0}{\log(2+k)} \quad (2.98)$$

Devido ao decaimento logarítmico da temperatura T_k , a estimação MAP proposta por Geman e Geman apresenta alta demanda computacional. Como alternativa a este método, Besag (1986) propõe o método ICM, baseado na maximização local da probabilidade condicional, como mostra a equação 2.99, através da escolha de x_s que maximize $P(X_s = x_s | \mathbf{y}, \partial_s)$ a cada ciclo. Este método, apesar de depender da configuração inicial de \mathbf{x} , apresenta rápida convergência, necessitando apenas de alguns ciclos do algoritmo 2.2.

$$P(X_s = x_s | \mathbf{y}, \partial_s) \propto P(Y_s = y_s | X_s = x_s) P(X_s = x_s | \partial_s) \quad (2.99)$$

Estimação MAP utilizando o método ICM

Seja \mathbf{x} uma realização pertencente ao espaço amostral $\Omega = \{1, 2, \dots, m\}^n$ e k denota o número de iterações necessárias para convergência.

```

1  escolha da modelagem do MRF utilizado na segmentação  $\mathbf{x}$ 
2   $x_i \leftarrow$  valoração aleatória //  $\forall x_i \in \mathbf{x}$ 
3
4  // atualização dos valores das variáveis aleatórias
5  para  $t=1$  até  $n$  faça
6      para  $v=1$  até  $m$  faça
7           $x_t \leftarrow \max\{P(Y_t = y_t | X_t = v) P(X_t = v | \partial_s)\}$ 
8
9  repita os passos 5–7 por  $k$  vezes
```

Algoritmo 2.2: Estimação MAP utilizando o método ICM.

O ICM maximiza a probabilidade condicional $P(X_s = x_s | y, \partial_s)$ durante cada iteração, sendo um algoritmo de relaxação determinístico de implementação simples. Tal método apresenta convergência rápida, necessitando de apenas algumas iterações e não permitindo que a probabilidade $P(x|y)$ diminua, pois é composta pela soma das energias locais, que estão sendo minimizadas (Manjunath et al. 1990).

Outra alternativa para estimação utilizando o MAP com SA é o método conhecido como maximização de probabilidades marginais, proposto por Marroquin et al. (1987). Este método tem como objetivo maximizar o número esperado de pixels classificados corretamente, minimizando a equação 2.100, onde x^* denota a classificação correta dos pixels, \hat{x} a realização estimada, $\delta(z) = 1$ se $z = 0$ e $\delta(z) = 0$ caso contrário.

$$\sum_{t=1}^n [1 - \delta(x_t^* - \hat{x}_t)] \quad (2.100)$$

A realização que minimiza o erro de classificação é aquela que maximiza a distribuição marginal posterior, desta maneira, o valor escolhido para o pixel \hat{x}_t é aquele que satisfaz a equação 2.101. Para que essa equação possa ser satisfeita, deve-se obter amostras de $P(x|y)$ por meio de um algoritmo iterativo de amostragem, armazenando cada realização obtida.

$$P(X_t = \hat{x}_t | y) \geq P(X_t = g | y), \quad \forall g \in \{1, 2, \dots, m\} \quad (2.101)$$

A determinação da probabilidade de uma realização da distribuição marginal de X_t é aproximada pela equação 2.102, onde k é o número de iterações necessárias para obter uma amostra estável de $P(x|y)$, N é um número grande o suficiente para obtenção de uma estimação razoável e que apresente custo computacional aceitável, sendo ambos parâmetros obtidos experimentalmente. Essa equação é utilizada pelo algoritmo 2.3, que apresenta os passos necessários para estimação da segmentação utilizando o método MPM.

$$\hat{P}(X_t = g | y) = \frac{1}{N - k} \sum_{i=k+1}^N \delta(x_t^i - g) \quad (2.102)$$

Estimação MAP utilizando o método MPM

Seja x uma realização pertencente ao espaço amostral $\Omega = \{1, 2, \dots, m\}^n$, $L = \{1, 2, \dots, m\}$ e $P(x|y)$ definido pela equação 2.95.

```

1  escolha da modelagem do MRF utilizado na segmentação x
2
3  // inicialização de x
```

```

4   $x_i \leftarrow g$  tal que maximize  $P(Y_s = y_s | X_t = g), \forall g \in L$ 
5
6  // obtenção das amostras da distribuição a posteriori  $P(x|y)$ 
7  para  $t=1$  até  $L$  faça
8      escolha  $g \in L$  aleatoriamente
9       $z_s \leftarrow g$ 
10      $z_r \leftarrow x_r \forall r \neq s$ 
11      $p \leftarrow \min \left\{ 1, \frac{P(z|y)}{P(x|y)} \right\}$ 
12     substitua  $x$  por  $z$  com probabilidade  $p$ .
13 repita os passos 7–12 por  $N$  vezes
14
15 // escolha da realização que minimiza erros de classificação
16 para  $t=1$  até  $n$  faça
17     atribua o valor  $\hat{x}_t$  ao pixel  $t$  que satisfaça
18      $\hat{P}(X_t = \hat{x}_t | y) > \hat{P}(X_t = g | y) \forall g \in L$ 

```

Algoritmo 2.3: Estimação MAP utilizando o método MPM.

2.4 Segmentação Baseada em Dependência Espacial

Esta seção descreve os principais métodos de segmentação de imagens baseados em dependência espacial modelada por campos aleatórios. Atenção principal é dada aos métodos que não utilizam informações obtidas antes do processo de segmentação. Sendo assim, o processo é dividido em duas etapas: a primeira, responsável pela obtenção de parâmetros para descrição das classes presentes na imagem e, a segunda, que utiliza esses parâmetros para efetuar a segmentação final.

Embora apenas a primeira etapa já possibilite a obtenção de uma segmentação, seus resultados são bastante grosseiros, desde que para se obter os parâmetros, a imagem é dividida em janelas e essas atribuídas às diferentes classes, o que causa uma grande falta de precisão e adaptação aos contornos finos presentes na imagem, como ilustra o resultado mostrado na figura 2.21(b). Com a segunda etapa, a segmentação inicial é ignorada total ou parcialmente, e uma segmentação baseada em pixels, ou segmentação fina, é efetuada, apresentando resultados mais acurados, conforme mostra a figura 2.21(c).

A indisponibilidade de informações previamente adquiridas torna a segmentação de imagens uma tarefa bastante complexa. A principal dificuldade está no fato que o modelo e seus parâmetros são desconhecidos e para calculá-los efetivamente seria necessário dispor da segmentação final, o que gera uma interdependência, pois o objetivo é justamente a segmentação.

Os trabalhos de Cohen e Cooper (1987) e Derin e Elliott (1987) apresentam a abordagem

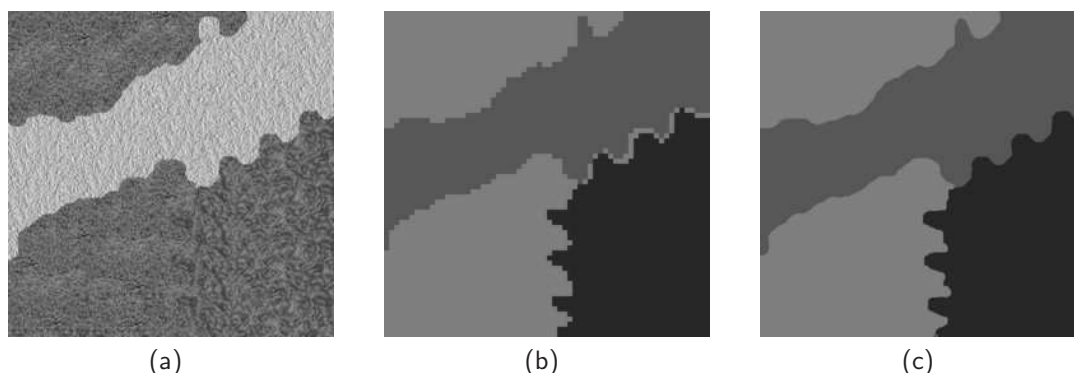


Figura 2.21: Segmentação de imagens. (a) imagem original; (b) segmentação obtida após a primeira etapa; (c) segmentação baseada em pixels.

que divide a segmentação baseada em dependência espacial em duas etapas. Na primeira, estima-se o número de classes distintas presentes na imagem bem como os parâmetros a elas associadas. Na segunda etapa, tais parâmetros são associados a um campo aleatório, que modela a distribuição dos rótulos das classes, quando algum algoritmo de relaxação é utilizado para aproximar a estimação MAP.

Para estimação dos parâmetros que descrevem cada classe presente na imagem, a maioria dos métodos de segmentação utiliza ou a extração de características por meio das abordagens estatística, geométrica e baseada em processamento de sinais, descritas na seção 2.2, ou a modelagem por meio de campos aleatórios de Markov (seção 2.2.4). Embora nem sempre seja possível a utilização dessa última alternativa, pois muitas vezes a textura não segue o modelo especificado.

Nessa abordagem de segmentação, a imagem de entrada é particionada em regiões com ou sem intersecção, a partir das quais são extraídas características, além de se estabelecer um MRF para modelar a distribuição dos rótulos das classes. Finalmente, essas duas informações são associadas por meio da formulação Bayesiana e a segmentação consiste na maximização da probabilidade *a posteriori*, ou na minimização de uma função de energia. A figura 2.22 ilustra o processo de segmentação de imagens baseado em dependência espacial.

O restante da seção apresenta os principais trabalhos que utilizam dependência espacial modelada por campos aleatórios de Markov para efetuar a segmentação de imagens coloridas e monocromáticas. Visando maior clareza, inicialmente é apresentada uma descrição geral dos métodos de segmentação e, em seguida, são abordadas especificamente a *obtenção de parâmetros* e a *segmentação baseada em pixels*, enfatizando principalmente os aspectos que apresentam maior interesse para o desenvolvimento deste trabalho.

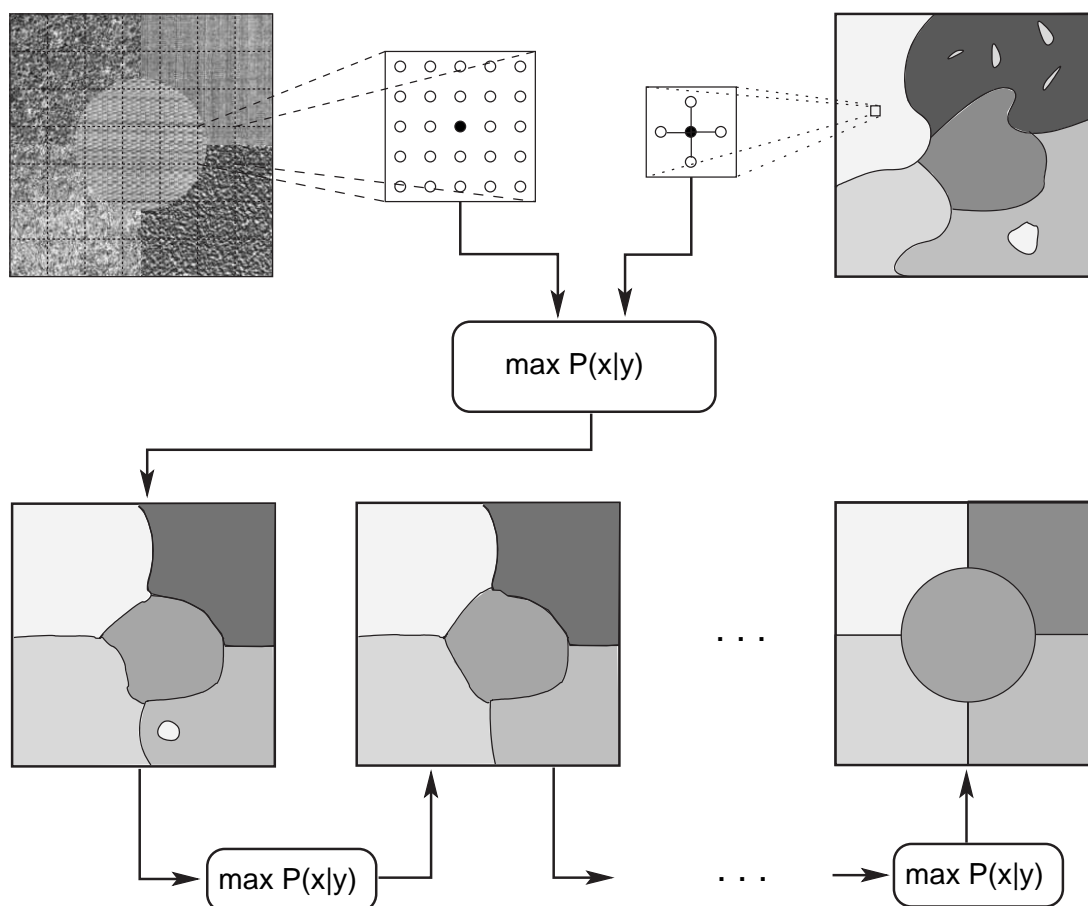


Figura 2.22: Processo de segmentação de imagens baseado em MRF e na formulação Bayesiana.

Descrição Geral dos Métodos

Simchony e Chellappa (1988) propõem um método que utiliza parâmetros previamente calculados, assumindo que a distribuição dos tons de cinza das texturas é modelada por um campo aleatório de Markov Gaussiano (GMRF, *Gaussian Markov random field*), definido por Besag (1974). Outro campo aleatório, este seguindo o modelo de Ising de segunda ordem, é utilizado para modelar a distribuição de probabilidade conjunta dos rótulos das classes. Os algoritmos ICM e SA são empregados com o objetivo de maximizar essa distribuição, tendo como condição inicial o resultado de uma segmentação obtida por meio da técnica de máxima verossimilhança (Theodoridis e Koutroumbas 2003).

Dentre os primeiros métodos de segmentação de imagens coloridas que utilizam conceitos de MRF para efetuar a segmentação fina estão os trabalhos de Daily (1989) e Huang et al. (1992). Além do relacionamento espacial retangular, normalmente utilizado, Daily efetua experimentos com relacionamentos hexagonais. Huang et al. utilizam uma segmentação inicial baseada em histogramas para as regiões suaves da imagem, enquanto uma segmentação fina, baseada em dependência espacial é aplicada nas fronteiras entre as regiões, utilizando-se o

ICM e o SA como técnicas de relaxação.

Efetuada algumas alterações no trabalho de Simchony e Chellappa, Manjunath e Chellappa (1991) propõem um método que assume que a distribuição de tons de cinza das texturas pode ser modelada por um GMRF de segunda ordem, cujos parâmetros são estimados a partir de janelas disjuntas, agrupadas por meio de uma medida de similaridade, resultando em uma segmentação inicial. Associando um MRF de segunda ordem, utilizado para modelar a distribuição conjunta dos rótulos das classes, com os parâmetros determinados inicialmente, a segmentação final é obtida por meio de algoritmos de relaxação.

Kervrann e Heitz (1995) empregam um procedimento de relaxação determinístico que simultaneamente efetua a segmentação e estima o número de classes, isto com a adição de um novo estado no MRF, o qual possibilita a criação de novas classes durante o processo de relaxação, embora para isso seja necessário a utilização de um limiar. Dado que uma medida de similaridade é empregada durante a relaxação, esse método não necessita de uma etapa inicial para obtenção de parâmetros.

Panjwani e Healey (1995) propõem um método para segmentação de imagens coloridas utilizando um GMRF com vizinhança de segunda ordem para modelar as interações entre os vizinhos. O método utiliza imagens coloridas no espaço cromático RGB, em que o vetor de características de cada região é composto pelos parâmetros do GMRF e pelas intensidades média nos canais R, G e B. A estimação dos parâmetros é efetuada de três maneiras distintas, pelo método *coding* (Besag 1974), pseudo-máxima verossimilhança e pela transformada de Fourier.

Fwu e Djuric (1996) propõem um método iterativo para simultaneamente, estimar os parâmetros do MRF e efetuar a segmentação da imagem. A cada ciclo do método, o número de classes presentes na imagem é incrementado utilizando-se uma estimação inicial para os parâmetros, que são melhorados através do ICM. Durante a segmentação, um critério para validação do agrupamento é utilizado para determinar o número de classes presentes na imagem.

Um método de segmentação de imagens que associa algoritmos genéticos com campo aleatório de Markov para efetuar a relaxação é proposto por Andrey e Tarroux (1998). Nesse trabalho, as texturas são consideradas como amostras do modelo generalizado de Ising e operadores genéticos são definidos para a etapa de relaxação. A associação entre algoritmos genéticos e técnicas de relaxação também é descrita por Tseng e Lai (1999), que propõem um método para segmentação de imagens multi-espectrais, empregando a mesma técnica apresentada por Huang et al. (1992) na segmentação inicial.

Kim et al. (1998) propõem um método de segmentação que, assim como nos trabalhos de Manjunath e Chellappa (1991) e Simchony e Chellappa (1988), modela as texturas por meio de um GMRF, entretanto, emprega-se o algoritmo EM (Dempster et al. 1977) na obtenção dos parâmetros do campo aleatório que modela cada textura, este algoritmo determina a qual distribuição de uma mistura pertence cada amostra.

Utilizando as definições propostas por Fwu e Djuric, Poggi e Ragozini (1999) apresentam um algoritmo para segmentação baseada em uma árvore binária, no qual, cada nodo apresenta um modelo de MRF. A imagem de entrada, associada à raiz da árvore, é segmentada em duas classes, criando dois novos nodos e, cada um desses nodos é segmentado em duas classes distintas. Esse processo é repetido até que uma condição de parada seja encontrada.

Embora Jackson e Landgrebe (2002) desenvolvam um classificador supervisionado, um MRF é aplicado com o algoritmo ICM para efetuar a classificação fina dos pixels, utilizando os vetores de características e os parâmetros obtidos a partir de um conjunto de treinamento. Com a introdução da dependência espacial no classificador, os resultados obtidos tornaram-se mais acurados.

Outro método de segmentação de imagens coloridas é proposto por Wesolkowski e Fieguth (2002). Tal método se baseia em um MRF de primeira ordem com conhecimento prévio do número de classes existentes e considera tanto relações globais quanto locais na minimização da função de energia. Para descrição individual de cada pixel, emprega-se a intensidade dos canais R, G e B com média zero, não utilizando portanto, uma região em torno do pixel para extração de características. Finalmente, o algoritmo baseado em SA é aplicado para efetuar a minimização da função de energia e atualizar as relações globais.

Uma versão estendida do método proposto por Poggi e Ragozini (1999) é apresentada por D'Elia et al. (2003), onde é criado o conceito de MRF estruturado por árvore (TS MRF, *Tree-Structured Markov Random Field*) para dirigir a segmentação da imagem utilizando uma árvore binária. Seja a árvore binária T , em que cada nodo interno t possui dois filhos $l(t)$ e $r(t)$, o conjunto de pixels a ser segmentado neste nodo é representado por $S^t \subset S$, onde $S = \{1, 2, \dots, n\}$. Por ser uma árvore binária, as variáveis aleatórias que compõem o MRF em cada nodo da árvore apresentam apenas dois valores possíveis, $l(t)$ ou $r(t)$.

Liapis et al. (2004) propõem um método de segmentação de imagens coloridas baseada em características de texturas extraídas a partir da transformada wavelet e histogramas de cores. Assumindo o número de classes como conhecido, uma segmentação inicial é efetuada com a eliminação das regiões heterogêneas e com o agrupamento daquelas que apresentam características semelhantes, isto determinado pela distância de *Bhattacharyya*. Finalmente, a

segmentação baseada em pixels é executada utilizando os algoritmos ICM, *fast marching level-set* (Sethian 1996) e funções de máxima verossimilhança, uma contendo informações sobre características de texturas e a outra com a adição de informações cromáticas. Os melhores resultados são obtidos com as duas primeiras abordagens.

A minimização da função de energia geralmente é efetuada sobre equações que apresentam dois componentes, onde o primeiro componente descreve as características dos pixels e o segundo modela a distribuição dos rótulos das classes. Deng e Clausi (2004) concentram-se na especificação de um fator, estimado no decorrer do processo de segmentação, para determinar a proporção que cada um desses componentes contribui para o sistema. Essa idéia também é a base do trabalho de Stewart et al. (2000), entretanto, sem a utilização direta de modelos Markovianos.

Obtenção de Parâmetros

Manjunath e Chellappa (1991) dividem a imagem em janelas disjuntas, para cada uma são estimados os parâmetros pelo método dos mínimos quadrados. O vetor de características é formado pelos quatro parâmetros do GMRF de segunda ordem e mais a média e a variância da janela, ou seja $\mathbf{F} = (\theta_1, \theta_2, \theta_3, \theta_4, \mu, \sigma^2)$. Para determinar quais janelas pertencem a uma mesma classe, utiliza-se a distância *Euclidiana normalizada*. Duas janelas i e j são agrupadas caso satisfaçam a equação 2.103, onde $d(\mathbf{F}^i, \mathbf{F}^j)$ denota a distância Euclidiana normalizada e ρ é um parâmetro definido experimentalmente. As regiões isoladas e aquelas que satisfazem a equação 2.103 para duas classes distintas são marcadas como ambíguas.

$$d(\mathbf{F}^i, \mathbf{F}^j) < \rho \max_{i,j} d(\mathbf{F}^i, \mathbf{F}^j) \quad (2.103)$$

Para determinar o número de classes presentes na imagem, Kim et al. (1998) propõem a medida mostrada na equação 2.104, onde m denota o número de classes, $L(t)$ o número de elementos contidos na classe t , representada por C_t , e $d(i, j)$ indica a distância entre os parâmetros dos elementos i e j . O valor de m que minimiza essa equação é aquele que apresenta o número de classes contidas na imagem.

$$D(M) = \frac{1}{m} \sum_{t=1}^m \left[\left(\frac{1}{L(t)} \sum_{i,j \in C_t} d(i, j) \right) - \left(\frac{1}{L(t)} \sum_{i,j \notin C_t} d(i, j) \right) \right] \quad (2.104)$$

No trabalho de Lau e Levine (2002), após a criação de um agrupamento hierárquico baseado em características de texturas e cores, são descritos três métodos para determinar

o número de classes presentes na imagem, isto por meio de índices da validação do agrupamento ou limiares. A abordagem que utiliza limiares, considerando a densidade e a distância dos agrupamentos i e j , baseia-se na equação 2.105, onde ρ e τ são limiares determinados experimentalmente. Quando o critério for satisfeito, ocorre a fusão dos agrupamentos i e j .

$$\text{densidade}(i, j) + \rho \cdot \text{distância}(i, j) > \tau \quad (2.105)$$

O *índice modificado de Hubert*, que considera que os agrupamentos assumem formas circulares, é a segunda abordagem apresentada por Lau e Levine. Para terceira abordagem é utilizada uma combinação de duas medidas não paramétricas que quantificam se os agrupamentos estão bem definidos. As medidas combinadas são a *norma NN*, que mede o isolamento de cada agrupamento e a *norma C*, que exhibe valores altos quando um dado agrupamento apresenta-se bem conectado.

Após a extração de características de regiões disjuntas da imagem, Liapis et al. (2004) utilizam a equação 2.106 como medida de homogeneidade para cada região, onde $d_B(W_0, W_i)$ denota a distância de Bhattacharyya entre a janela que está sendo analisada e um de seus oito vizinhos mais próximos.

$$M(W_0) = \frac{1}{8} \sum_{i=1}^8 d_B(W_0, W_i) - \min_{i=1, \dots, 8} d_B(W_0, W_i) \quad (2.106)$$

Para determinar quais regiões são heterogêneas, Liapis et al. criam uma ordenação segundo $M(W_i)$, gerando desta maneira, um histograma bimodal. Finalmente, a equação 2.107 é utilizada para determinar onde ocorre a mudança dentro do histograma. Tal equação é uma aproximação da segunda derivada calculada sobre a distribuição do histograma, onde $n+1$ e $n-1$ indicam os vizinhos da função de homogeneidade para o bloco n , segundo a ordenação efetuada.

$$M(W_{n+1}) - 2M(W_n) + M(W_{n-1}) \quad (2.107)$$

Com o descarte das regiões heterogêneas, um método hierárquico aglomerativo de agrupamento é aplicado, unindo duas janelas vizinhas que apresentarem distância menor que o valor limiar T_p , mostrado na equação 2.108, até que reste apenas o número de classes previamente determinado. Nessa equação, n indica o número de regiões homogêneas e W_0 denota a região que está sendo atualmente testada.

$$T_p = \frac{1}{2} \max_{i=1, \dots, n} \left\{ \frac{1}{n} \sum_{i=1}^n d_B(W_0, W_i) \right\} + \min_{i=1, \dots, n} \left\{ \frac{1}{n} \sum_{i=1}^n d_B(W_0, W_i) \right\} \quad (2.108)$$

Segmentação Baseada em Pixels

Manjunath e Chellappa (1991) utilizam a relaxação baseada em duas abordagens para obter a realização do MRF que indica a segmentação final, uma determinística e outra estocástica. A primeira utiliza um algoritmo semelhante ao ICM para minimizar equação 2.109, onde N_s denota a vizinhança de segunda ordem de s , V_{sl} assume 1 quando o pixel s pertence a classe l e 0 caso contrário, β representa a interação entre regiões vizinhas e $U(s, l)$, modelado como um GMRF, apresenta informações sobre os parâmetros que caracterizam o pixel s .

$$E = \frac{1}{2} \sum_s \sum_l U(s, l) V_{sl} - \frac{\beta}{2} \sum_s \sum_l \sum_{s' \in N_s} V_{s'l} V_{sl} \quad (2.109)$$

O algoritmo estocástico utilizado por Manjunath e Chellappa se baseia na minimização da classificação incorreta, por meio do MPM. Testes são efetuados combinando a relaxação determinística com o aprendizado estocástico. Essa combinação apresenta a vantagem de requerer um número menor de ciclos que os algoritmos baseados em resfriamento simulado e a segmentação resultante é melhor que a obtida apenas com a relaxação determinística.

Com o acréscimo de mais um estado, ρ , no MRF que modela a distribuição dos rótulos das classes, o método de Kervrann e Heitz (1995) não necessita da etapa de estimação dos parâmetros, pois o número de classes pode ser determinado durante a estimação MAP, entretanto, torna-se necessário a utilização de um limiar para determinar quando uma nova classe deve ser criada. Dessa maneira, o objetivo de Kervrann e Heitz é obter o número de classes R e estimar a realização do campo aleatório \mathbf{x} , isto através da minimização da função de energia mostrada na equação 2.110, onde U_1 e U_2 dependem, respectivamente, das características das regiões e do MRF que modela a distribuição dos rótulos.

$$\hat{\mathbf{x}} = \min_{\mathbf{x} \in \Omega} \{U_1(\mathbf{x}, \mathbf{o}, R) + U_2(\mathbf{x})\} \quad (2.110)$$

A função de potencial U_2 considera vizinhança de segunda ordem para modelar a distribuição dos rótulos das classes, enquanto o termo U_1 é composto basicamente pela distância *Kolmogorov-Smirnov* entre as características do pixel visitado e as características das classes. O rótulo ρ é atribuído a um pixel se a distância de Kolmogorov-Smirnov entre a distribuição

das características do pixel visitado e a distribuição global de todas as classes existentes for maior do que um limiar.

O ICM é utilizado para minimização local da equação 2.110. Quando a realização apresenta valores estáveis, os componentes conexos com o rótulo ρ são atribuídos a novas classes e o ICM é reiniciado, agora considerando essas novas classes. Esse processo é repetido até que a segmentação se estabilize por completo.

Jackson e Landgrebe (2002) empregam o ICM para executar a segmentação fina dos pixels, utilizando o conjunto de treinamento para determinação do vetor de médias e a matriz de covariância para todas as classes. Como descrito na seção 2.3, o ICM maximiza a equação 2.111, para isso são definidas as equações 2.112 e 2.113.

$$P(X_s = x_s | y, \partial_s) \propto P(Y_s = y_s | X_s = x_s) P(X_s = x_s | \partial_s) \quad (2.111)$$

Nas equações 2.112 e 2.113, l denota o número de componentes no vetor de características y_s , Z representa a função de partição, μ_{x_s} e Σ_{x_s} representam o vetor de médias e a matriz de covariância das características para a classe x_s , onde $x_s \in \{1, 2, \dots, m\}$. Na segunda equação, β é um parâmetro determinado experimentalmente, utilizado para definição da função de potencial (definição 2.9) e C_s representa o conjunto de cliques que contém o pixel s .

$$P(Y_s = y_s | X_s = x_s) = \frac{1}{(2\pi)^{l/2} |\Sigma_{x_s}|^{1/2}} \exp \left[-\frac{1}{2} \{ (y_s - \mu_{x_s})^T \Sigma_{x_s}^{-1} (y_s - \mu_{x_s}) \} \right] \quad (2.112)$$

$$P(X_s = x_s | \partial_s) = \frac{1}{Z} \exp \left\{ -\beta \sum_{t \in C_s} [1 - \delta(x_s - x_t)] \right\} \quad (2.113)$$

A cada ciclo do algoritmo ICM, efetua-se a minimização da função de energia local mostrada na equação 2.114, a qual é obtida pelo desenvolvimento das equações anteriores. Os parâmetros utilizados são obtidos na etapa de treinamento, c é uma constante que independe da atribuição do pixel para alguma classe e k denota o número de ocorrências de classes distintas de x_s em cliques que contêm s .

$$\hat{x}_s = \min_{1 \leq r \leq L} [\ln |\Sigma_r| + (y_s - \mu_r)^T \Sigma_r^{-1} (y_s - \mu_r) + 2k\beta + c] \quad (2.114)$$

Deng e Clausi (2004) acrescentam o fator de ponderação $\alpha(t)$ na função de energia a ser minimizada, como mostra a equação 2.115, onde E_R depende do MRF que modela a distribuição dos rótulos das classes e E_F depende das características extraídas dos pixels.

$$E = E_R + \alpha(t)E_F \quad \alpha(t) = c_1 0.9^t + c_2 \quad (2.115)$$

O termo $\alpha(t)$, mostrado na equação 2.115, tem seu valor alterado no decorrer do processo de relaxação de modo a alterar a influência da função de energia E_F . Em seus experimentos, Deng e Clausi utilizam $c_1 = 80$ e $c_2 = 1/m$, onde m representa o número de classes presentes na imagem.

CAPÍTULO 3

METODOLOGIA PROPOSTA

Este capítulo descreve a metodologia proposta para o desenvolvimento de um segmentador não supervisionado aplicável a imagens coloridas e monocromáticas, para isso, conforme as considerações feitas na seção 2.4, o método é dividido em duas etapas, as quais possibilitam a obtenção de uma segmentação que apresente elevada qualidade.

A primeira etapa inclui a determinação do número de classes presentes na imagem bem como a estimação dos parâmetros que as descrevem. Na segunda etapa, tais parâmetros são associados a um MRF que modela a distribuição de rótulos das classes com objetivo de criar uma função de energia a ser minimizada posteriormente por meio do algoritmo ICM. A segmentação final é obtida diretamente da realização apresentada pelo MRF ao final do ICM.

Para que a descrição do método seja efetuada, este capítulo é dividido em três seções, a primeira descreve aspectos gerais do modelo de segmentação adotado, ressaltando algumas características relevantes. As duas últimas seções estão relacionadas aos aspectos mais técnicos envolvidos no desenvolvimento da metodologia proposta, onde são descritas em detalhes a primeira e segunda etapas do segmentador.

3.1 Aspectos Gerais

Em algumas áreas de aplicação a aquisição prévia de informações apresenta-se como uma tarefa intrincada ou até mesmo inviável, tornando-se necessária a criação de metodologias que extraíam informações a partir da própria imagem a ser segmentada, o que torna o desenvolvimento de um método de segmentação uma tarefa complexa, pois quaisquer parâmetros necessários devem ser obtidos durante a execução do processo.

Conforme descrito na seção 2.4, os métodos que não utilizam dados previamente calculados geralmente dividem o processo de segmentação em duas etapas. A primeira é responsável pela obtenção de parâmetros e a segunda efetua a segmentação propriamente dita. Portanto, com o objetivo de promover a segmentação não supervisionada, ou seja, sem conhecimento prévio do número de classes e nem dos parâmetros que as descrevem, a metodologia proposta neste trabalho divide-se em duas etapas.

Alguns aspectos podem ser ressaltados durante a primeira etapa do método proposto,

aquela que possibilita que os parâmetros sejam obtidos a partir de classes homogêneas. Dentre eles, a característica de ser aplicável tanto em imagens monocromáticas quanto em coloridas; a dispensabilidade da utilização da modelagem para as texturas; as regiões heterogêneas presentes na imagem são detectadas de maneira simples; com uma pequena alteração na inicialização do método de agrupamento *k-means* são obtidos agrupamentos estáveis; os parâmetros utilizados durante segunda etapa são extraídos a partir de classes essencialmente homogêneas; e, finalmente, a utilização de critérios simples para efetuar a validação do agrupamento.

Muitos métodos de segmentação tentam modelar as texturas a partir de campos aleatórios, tarefa que necessita de estimadores para obtenção dos parâmetros, o que não se apresenta como um processo simples. Adicionalmente, mesmo sendo efetuada a estimação dos parâmetros, algumas texturas não seguem a distribuição descrita pelo modelo utilizado. Para não depender de modelos específicos, embora testes sejam efetuados com um modelo de MRF, são utilizados em sua maioria métodos estatísticos, geométricos e baseados em processamento de sinais para obtenção dos vetores de características.

Dado que as medidas utilizadas para descrever as classes e as regiões são obtidas a partir de métodos de análise de texturas, a metodologia proposta pode ser aplicada a imagens coloridas ou monocromáticas apenas com a alteração do método empregado para efetuar a extração de características. A seção 4.2 apresenta resultados da segmentação para ambas as classes de imagens.

O método de agrupamento *k-means*, baseado na minimização de uma função de custo, apresenta baixo custo computacional, entretanto, não utilizado algumas vezes por apresentar sensibilidade à escolha inicial dos centros para os agrupamentos. Uma alteração efetuada nessas condições iniciais proporciona a obtenção de resultados estáveis, permitindo que as classes homogêneas sejam detectadas de maneira precisa.

Para que os parâmetros utilizados na descrição de uma determinada classe sejam representativos, é de fundamental importância que sejam extraídos a partir de regiões que pertençam a uma mesma classe, caso contrário, podem ocorrer instabilidades que acarretam uma segmentação final imprópria. Para isso é proposto um método simples para detecção de classes homogêneas, este caracterizado por não utilizar valores limiares para efetuar tal detecção, necessitando apenas do agrupamento resultante do *k-means*.

A determinação do número de classes contidas na imagem é uma etapa crucial do método, pois sem o cômputo correto dessa informação todo o processo é comprometido. Muitas vezes, valores limiares são utilizados, tornando-a específica para cada aplicação. Este trabalho avalia alguns critérios utilizados normalmente para a seleção de características, porém, os resultados

apresentados na seção 4.2 demonstram sua aplicabilidade na validação do agrupamento.

Na segunda etapa, o principal aspecto a ser ressaltado está no fato de se utilizar a dependência espacial, entretanto, outros fatores relevantes devem ser enfatizados na metodologia proposta. Dentre eles, a flexibilidade na definição da função de energia a ser minimizada, algumas características do ICM, tais como baixa demanda computacional, insensibilidade quanto ao fenômeno denominado transição de fase (Besag 1986), a dispensabilidade de aproximações para função de partição, e os bons resultados por ele proporcionados (Dubes e Jain 1989, Liapis et al. 2004).

Conforme descrito em outros capítulos, a segmentação baseada em dependência espacial proporciona resultados superiores aqueles obtidos apenas com a utilização de métodos de segmentação tradicionais. Em sua modelagem utiliza-se um MRF para descrever a distribuição dos rótulos das classes. Ao final da execução, a realização desse MRF contém de maneira direta a segmentação que aproxima a estimação MAP descrita na seção 2.3.

Atendendo a algumas restrições, a probabilidade *a posteriori* é modelada por um MRF que segue a distribuição apresentada na seção 2.3, desta maneira, a estimação MAP é transformada na minimização de uma função de energia, a qual pode ser escolhida de maneira a compreender as restrições apresentadas para cada área de aplicação. Este trabalho utiliza uma função de energia determinada pela combinação da distribuição Gaussiana multivariada com um MRF que considera apenas cliques compostos por dois elementos.

Visando à minimização da função de energia emprega-se o algoritmo ICM, que apresenta menor demanda computacional que outros métodos de relaxação (Simchony e Chellappa 1988), embora tenda a convergir para regiões de mínimos locais, pois não permite a obtenção de estados com maior energia. Este trabalho apresenta uma alteração no ICM para que estados de maior energia sejam permitidos. Os resultados mostrados na seção 4.2 indicam que a energia é minimizada com maior intensidade quando essa alteração é aplicada.

3.2 Segmentação Inicial

O objetivo desta etapa é a determinação do número de classes e a estimação dos parâmetros que aproximem a distribuição de características por elas seguida. Para tal, deve-se inicialmente determinar a localização das classes homogêneas presentes na imagem, caso contrário, pode haver a caracterização de regiões contidas nas fronteiras entre as classes, ocasionando imprecisão na segmentação final. O diagrama da figura 3.1 apresenta os passos efetuados durante a segmentação inicial da imagem.

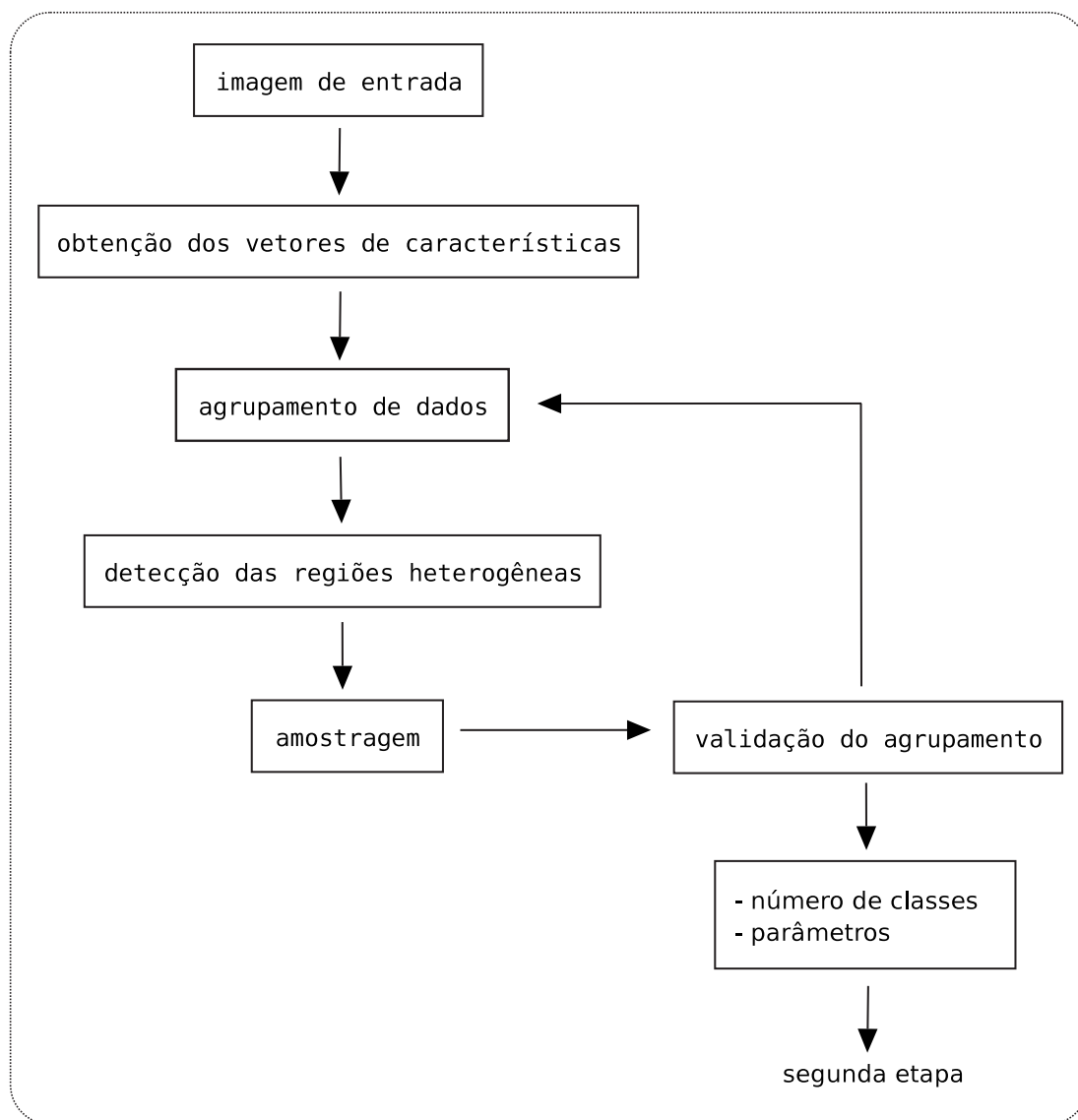


Figura 3.1: Diagrama ilustrando a segmentação inicial. Após a validação do agrupamento tem-se como resultado o número de classes e os parâmetros que as descrevem.

Obtenção dos Vetores de Características

Desde que o método proposto não emprega um conjunto de treinamento, a caracterização de cada classe deve ser efetuada utilizando apenas informações contidas na imagem a ser segmentada. Dessa maneira, para que os parâmetros necessários sejam obtidos, regiões de interesse devem ser devidamente identificadas.

Se o resultado da localização das regiões não for satisfatório, a segmentação final será prejudicada, pois a aproximação da estimação MAP, efetuada durante a segunda etapa, associa o modelo da distribuição dos rótulos aos parâmetros que descrevem cada classe, estes estimados a partir das referidas regiões. Sendo assim, o resultado obtido é uma segmentação imprópria, inviabilizada pela minimização de uma função de energia mal definida.

Tendo sido mencionada a importância de uma amostragem inicial de boa qualidade, propõe-se que, ao contrário da maioria dos métodos encontrados na literatura, que a imagem seja dividida em R regiões composta de $M \times M$ pixels, de modo que a intersecção entre regiões distintas seja permitida. Dessa maneira, pode-se determinar quais pixels estão situados em regiões heterogêneas e eliminar a possibilidade de que sejam utilizados na estimação dos parâmetros.

Para cada região amostrada, características de texturas são extraídas, através dos métodos de análise de texturas descritos na seção 2.2, de modo que cada medida retornada seja acrescentada em um vetor de características, utilizado posteriormente na descrição dessa região. Ou seja, ao invés de utilizar os níveis de cinza dos pixels, calcula-se medidas que resumem as propriedades de cada região, proporcionando, desta maneira, a redução da dimensionalidade dos dados, entretanto, sem perda significativa na capacidade discriminatória.

Por exemplo, considerando regiões compostas de 16×16 pixels, se o vetor de características para cada região fosse preenchido apenas com o tom de cinza dos pixels, resultaria em um vetor composto por 256 componentes altamente correlacionados, inclusive acarretando aumento no custo computacional ao se determinar medidas de similaridade entre dois vetores quaisquer. No entanto, com a aplicação de métodos de análise de texturas um número reduzido de medidas é utilizado.

Definido o método de análise de texturas a ser utilizado, efetua-se sua aplicação para cada uma das R regiões da imagem, já considerando aquelas que possuem intersecção. O vetor de características é definido na equação 3.1, onde l denota o número de medidas utilizadas e $i = 1, 2, \dots, R$. Deve-se observar que y_i pode ser composto por medidas resultantes da combinação de métodos de análise de texturas.

$$y_i = \{y_{i1}, y_{i2}, \dots, y_{il}\} \quad (3.1)$$

Uma desvantagem em utilizar diretamente o vetor de características para o cálculo de distâncias se deve às diferenças de escalas encontradas nas medidas, ocasionando ponderações indevidas. Para minimizar este problema, efetua-se a padronização dessas medidas por meio da equação 3.2, onde μ e σ denotam os vetores de média e variância, obtidos pelas equações 3.3 e 3.4, respectivamente.

$$z_i = \frac{y_i - \mu}{\sigma} \quad (3.2)$$

$$\mu = \frac{1}{R} \sum_{i=1}^R y_i \quad (3.3)$$

$$\sigma^2 = \frac{1}{R-1} \sum_{i=1}^R (y_i - \mu)^2 \quad (3.4)$$

Outra abordagem que permite que a dimensionalidade dos dados seja reduzida é denominada análise de componentes principais (PCA, *principal component analysis*). Entretanto, é necessário que seja utilizada em conjunto com algum método de seleção de características, pois a PCA efetua uma transformação linear em que medidas distintas passam a apresentar correlação nula, porém não atua na separabilidade entre classes, como ilustrado na figura 3.2. Desse modo, a PCA não é considerada durante o método de segmentação, embora seja utilizada na avaliação dos métodos de análise de texturas.

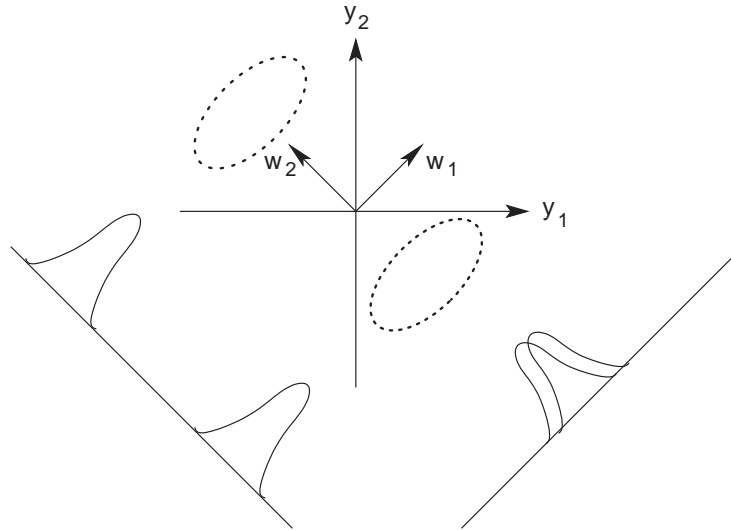


Figura 3.2: Aplicação da PCA, acarretando sobreposição entre as distribuições das medidas se escolhido o componente associado ao maior autovalor, neste caso w_1 .

A equação 3.5 apresenta a transformação linear efetuada pela PCA, onde G denota a matriz dos autovetores obtidos a partir da matriz de covariância definida em 3.6. Dado que os elementos da diagonal principal da matriz de autovalores representam a proporção da variância explicada pelas variáveis, a redução da dimensionalidade é efetuada com a manutenção das variáveis correspondentes aos autovalores que apresentam maior valor.

$$w_i = G y_i \quad (3.5)$$

$$\Sigma_i = \begin{pmatrix} \sigma_1\sigma_1 & \sigma_1\sigma_2 & \dots & \sigma_1\sigma_l \\ \sigma_2\sigma_1 & \sigma_2\sigma_2 & \dots & \sigma_2\sigma_l \\ \vdots & & & \vdots \\ \sigma_l\sigma_1 & \sigma_l\sigma_2 & \dots & \sigma_l\sigma_l \end{pmatrix} \quad (3.6)$$

Neste ponto da execução do método, encontram-se disponíveis os vetores de características compostos por medidas padronizadas. Esses vetores serão utilizados na determinação de um agrupamento inicial, a partir do qual será criado um histograma tridimensional que auxilia na determinação das regiões da imagem que apresentam características homogêneas.

Agrupamento de Dados

Para que sejam obtidas as classes homogêneas, primeiro os dados são agrupados de modo que as regiões atribuídas a cada classe apresentem características semelhantes. Deve-se observar que apenas com esse agrupamento pode-se efetuar a estimação dos parâmetros, entretanto, esta abordagem não obteria parâmetros representativos pois regiões heterogêneas ainda estão sendo consideradas, como ilustrado na figura 3.3. Portanto, um refinamento deve ser efetuado, o qual é proporcionado neste trabalho pela utilização de um histograma calculado a partir do agrupamento resultante da presente etapa.

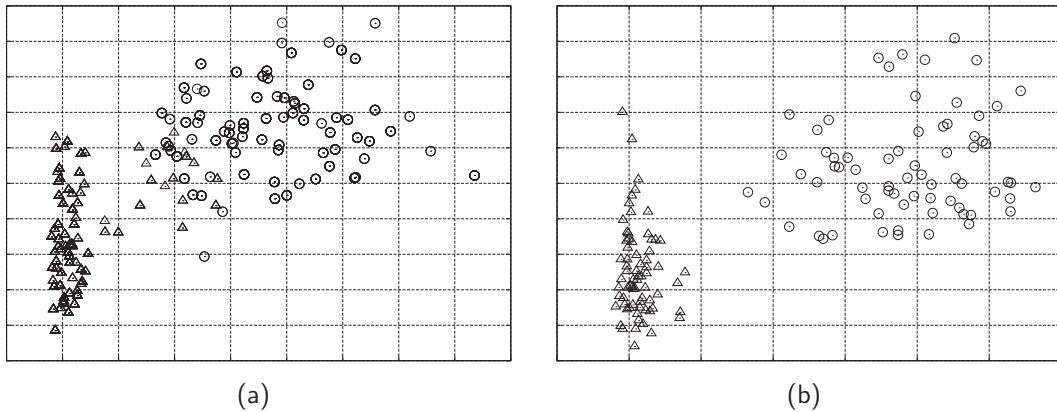


Figura 3.3: Espaço de características bidimensional. (a) espaço após o agrupamento de dados, ainda considerando as regiões heterogêneas; (b) espaço após o refinamento proporcionado pelo histograma.

Embora diversas metodologias têm sido empregadas para determinar quais regiões apresentam características semelhantes (seção 2.4), este trabalho se baseia naquela que visa à minimização da função de custo mostrada na equação 3.7. Nessa equação, θ denota a matriz de parâmetros, m o número de classes e U uma matriz $R \times m$ com $u_{ij} = 1$ se a i -ésima amostra pertencer à j -ésima classe e $u_{ij} = 0$ caso contrário.

$$J(\theta, U) = \sum_{i=1}^R \sum_{j=1}^m u_{ij} \|z_i - \theta_j\|^2 \quad (3.7)$$

Para minimização da equação 3.7 é apresentada uma modificação do conhecido algoritmo *k-means* (Duda e Hart 1973). O algoritmo original é dependente da configuração inicial, ou seja, o agrupamento final está relacionado com a escolha inicial dos representantes de cada classe. Para reduzir esse problema, propõe-se que o *k-means* seja executado algumas vezes, empregando configurações iniciais distintas para cada uma e, ao final, escolhe-se o agrupamento que minimiza o valor da função de custo. Os passos desse método são apresentados no algoritmo 3.1.

Agrupamento de dados

Agrupamento efetuado com o algoritmo *k-means*. $\text{random}(i)$ representa uma função que retorna amostras entre 1 e i da distribuição de probabilidade uniforme. z_i denota o vetor de características da i -ésima região, apresentando medidas padronizadas, onde $i = 1, 2, \dots, R$. N_a representa o número de agrupamentos a serem obtidos. A alteração efetuada é mostrada nas linhas 1 e 26, o algoritmo original retorna o primeiro agrupamento obtido, enquanto este retorna aquele que minimiza o valor de $J(\theta, U)$.

```

1  para  $i = 1$  até  $N_a$  faça
2      // inicialização dos centros dos agrupamentos
3      para  $j = 1$  até  $m$  faça
4           $r \leftarrow \text{random}(R)$ 
5           $\theta_j \leftarrow z_r$ 
6
7      enquanto  $J_i^t(\theta, U) \neq J_i^{t-1}(\theta, U)$  faça
8          // atribuição ao centro mais próximo
9          para  $j = 1$  até  $R$  faça
10              $t \mid \min_{1 \leq t \leq m} \|z_j - \theta_t\|^2$ 
11             para  $l = 1$  até  $m$  faça
12                 se  $l = t$  então
13                      $u_{il} \leftarrow 1$ 
14                 senão
15                      $u_{il} \leftarrow 0$ 
16
17             // atualização dos centros
18              $\theta_c \leftarrow 0$  para  $c = 1, 2, \dots, m$ 
19             para  $j = 1$  até  $R$  faça
20                 para  $l = 1$  até  $m$  faça
21                     se  $u_{jl} = 1$  então
22                          $c \leftarrow l$ 
23                      $\theta_c \leftarrow \theta_c + z_j/R$ 
24
25     // determinação de qual agrupamento deve ser retornado

```

26 **retorna** $U \mid \min_{1 \leq c \leq N_a} J_c(\theta, U)$

Algoritmo 3.1: Agrupamento de dados.

Devido ao seu baixo custo computacional, a modificação apresentada não influencia significativamente o tempo de execução do método de segmentação proposto, entretanto, os resultados obtidos apresentam-se mais estáveis. Sendo assim, se o algoritmo for executado diversas vezes sobre uma mesma entrada, o agrupamento resultante tende a permanecer estável, o que não ocorre com o *k-means* original (alguns experimentos são mostrados na seção 4.2).

Também foram efetuados testes com uma modificação do algoritmo *k-means* que utiliza SA para a minimização de $J(\theta, U)$. Embora essa abordagem garanta a convergência para o mínimo global, o tempo computacional torna-se inaceitável (Kirkpatrick et al. 1983), apresentando o mesmo problema do algoritmo de estimação MAP utilizando SA (seção 2.3).

Detecção de Regiões Heterogêneas e Amostragem

Como descrito anteriormente, é necessário que o agrupamento obtido seja refinado de modo que cada classe contenha apenas regiões homogêneas, caso contrário, os parâmetros podem ser estimados de maneira inadequada. Sendo assim, as regiões heterogêneas devem ser detectadas e excluídas do agrupamento obtido com o algoritmo *k-means*.

A amostragem inicial que permite a sobreposição entre as regiões possibilita que essa detecção seja efetuada de maneira simplificada, por meio da criação de um histograma tridimensional que tabula a quais classes cada pixel foi atribuído durante o agrupamento de dados. Os pixels que tenham sido atribuídos a mais de uma classe durante a criação do agrupamento de dados são considerados como pertencentes a regiões heterogêneas da imagem.

O algoritmo 3.2 descreve os passos necessários para criação desse histograma e para a determinação das regiões heterogêneas. Ao final de sua execução, as regiões que apresentem características homogêneas são amostradas para que seja efetuada tanto a validação do agrupamento quanto a estimação dos parâmetros.

Histograma tridimensional

Criação do histograma e determinação das regiões heterogêneas. R denota o número de regiões amostradas com sobreposição e submetidas ao agrupamento de dados, r_i representa uma dessas regiões, onde $i = 1, 2, \dots, R$. $h(x, y, c)$ representa o histograma

tridimensional, onde x e y são as coordenadas do pixel localizado em (x, y) e c denota a classe a ele atribuída, ou seja, $c = 1, 2, \dots, m$.

```

1 // criação do histograma
2  $\forall x, y, c \ h(x, y, c) = 0$ 
3 para  $i = 1$  até  $R$  faça
4    $c \leftarrow$  classe atribuída a  $r_i$ 
5   para todo pixel com coordenadas  $(x, y) \in r_i$  faça
6      $h(x, y, c) \leftarrow h(x, y, c) + 1$ 
7
8 // determinação das regiões heterogêneas
9 para todas as coordenadas  $(x, y)$  da imagem faça
10   se  $h(x, y, f) \neq 0$  e  $h(x, y, g) \neq 0 \ \forall f, g \in \{1, 2, \dots, m\}$  e  $f \neq g$ 
11     pixel em  $(x, y)$  pertence a uma região heterogênea
12   senão
13     pixel em  $(x, y)$  pertence a uma região homogênea

```

Algoritmo 3.2: Histograma tridimensional.

O processo de amostragem das regiões contidas em classes homogêneas dá-se de maneira aleatória, ou seja, qualquer região pode ser amostrada, desde que todos os pixels nela contidos pertençam à mesma classe, atingindo a condição de parada quando um certo número de amostras estiverem disponíveis para cada classe. O tamanho das janelas obtidas na amostragem deve ser o mesmo daquelas utilizadas posteriormente pelo ICM, pois os parâmetros aqui obtidos são utilizados durante sua execução (alguns experimentos apresentando variação no tamanho da janela são efetuados na seção 4.1). A figura 3.4 ilustra o processo descrito.

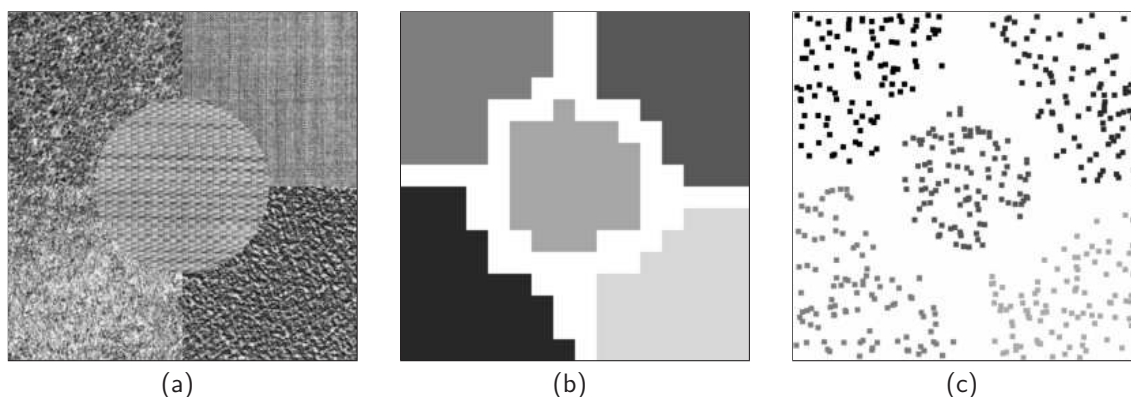


Figura 3.4: Amostragem das classes homogêneas. (a) imagem original; (b) regiões obtidas a partir do histograma; (c) regiões amostradas, utilizando 100 amostras para cada classe.

Para que a sobreposição necessária à criação do histograma tridimensional não apresente custo significativo, o deslocamento utilizado na amostragem inicial é sempre $M/2$, ou seja, metade do tamanho da região amostrada. Dessa maneira, para cada pixel, desconsiderando aqueles localizados nas bordas da imagem, há quatro entradas no histograma.

Validação do Agrupamento

A validação do agrupamento, responsável pela determinação do número de classes presentes na imagem, procura analisar se os agrupamentos apresentam-se bem separados no espaço de características. Para isso, são estimados alguns parâmetros a partir das classes homogêneas resultantes do histograma.

Embora existam outras abordagens utilizadas para determinação do número de classes, como descrito na seção 2.4, este trabalho avalia três critérios para determinar a separabilidade entre as classes. Tais critérios são baseados em matrizes de espalhamento (SM, *scatter matrices*), normalmente aplicadas na seleção de características (Theodoridis e Koutroumbas 2003).

As matrizes de espalhamento são mostradas nas equações 3.8 a 3.10, sendo denominadas como matriz de espalhamento intra-classe, matriz de espalhamento inter-classes e matriz de mistura de espalhamento, respectivamente. Tais matrizes fornecem a base para criação dos critérios utilizados na determinação do número de classes presentes na imagem. Os parâmetros P_i , S_i e μ_i denotam, respectivamente, a probabilidade *a priori*, a matriz de covariância e o vetor de médias da i -ésima classe, onde P_i é determinada conforme a área que cada classe representa após a eliminação das regiões heterogêneas.

$$S_w = \sum_{i=1}^m P_i S_i \quad (3.8)$$

$$S_b = \sum_{i=1}^m P_i (\mu_i - \mu_0)(\mu_i - \mu_0)^T \quad (3.9)$$

$$S_m = S_w + S_b \quad (3.10)$$

Finalmente, o parâmetro μ_0 representa a média ponderada total entre as classes homogêneas, calculada conforme a equação 3.11.

$$\mu_0 = \sum_{i=1}^m P_i \mu_i \quad (3.11)$$

Após a definição das matrizes de espalhamento, são apresentados os critérios que determinam o grau de separabilidade entre as classes no espaço de características. Essa informação se apresenta como um bom indicativo do número de classes, pois se este for sub-estimado, agrupamentos que apresentam centros distantes uns dos outros serão atribuídos a uma mesma classe e, caso ocorra a super-estimação, um único agrupamento será particionado em duas classes apresentando centros muito próximos.

Tendo em vista a separabilidade entre classes distintas e a compacidade presente em cada

agrupamento, o primeiro critério, mostrado na equação 3.12, apresenta valores altos quando as amostras estão bem agrupadas em torno da média e agrupamentos distintos apresentam-se bem separados, indicando a existência de uma boa separabilidade entre os dados (Theodoridis e Koutroumbas 2003).

$$J_1 = \frac{\text{traço}\{S_m\}}{\text{traço}\{S_w\}} \quad (3.12)$$

Os outros dois critérios, J_2 (equação 3.13) e J_3 (equação 3.14), apresentam a vantagem, em relação ao primeiro, de serem invariantes quanto a transformações lineares (Theodoridis e Koutroumbas 2003), nas equações 3.13 e 3.14, $|S_m|$ denota o determinante da matriz S_m . Entretanto, o critério J_2 apenas é válido para matrizes de espalhamento que sejam positivas definidas.

$$J_2 = \frac{|S_m|}{|S_w|} \quad (3.13)$$

$$J_3 = \text{traço}\{S_w^{-1}S_m\} \quad (3.14)$$

Com o objetivo de avaliar se uma matriz simétrica é positiva definida, pode-se utilizar o determinante das matrizes definidas em 3.15, onde A denota uma matriz $n \times n$, $A(n) = A$ e $k = 1, \dots, n$. A matriz A é positiva definida se $\det A(k) > 0$ (Hohn 1958).

$$A(k) = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} \quad (3.15)$$

Um exemplo dos valores assumidos em cada um dos critérios é apresentado no gráfico da figura 3.5, onde a abscissa indica o número de classes e a ordenada o valor resultante para cada critério, tais valores são obtidos com aplicação da primeira etapa do método de segmentação proposto, considerando um número distinto de classes. Outros resultados e discussões sobre esses critérios são apresentados na seção 4.2.

Após o número de classes ter sido determinado, sendo aquele onde os critérios de separabilidade entre as classes apresentam valor máximo, efetua-se a estimação da matriz de covariância e do vetor de médias para cada classe presente na imagem, a partir das mesmas regiões amostradas para validação do agrupamento. Tais parâmetros serão utilizados na definição da função de energia a ser minimizada pelo ICM.

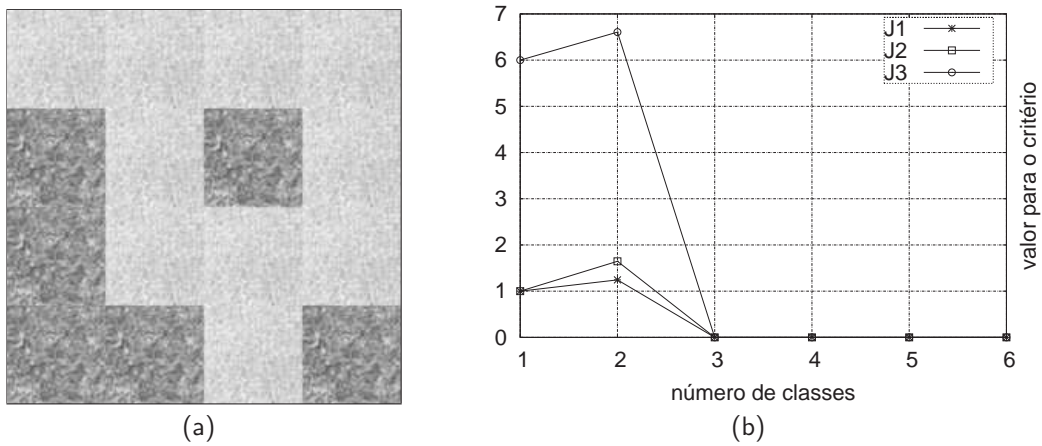


Figura 3.5: Validação do agrupamento. (a) imagem original; (b) valores resultantes para cada um dos critérios.

3.3 Segmentação Baseada em Pixels

Após a estimação dos parâmetros a partir das classes homogêneas e a determinação do número de classes, inicia-se a segmentação baseada em pixels, composta pelos passos apresentados no diagrama da figura 3.6. A execução de um algoritmo de relaxação, visando à minimização de uma função de energia definida pela combinação desses parâmetros com um MRF, permite que seja obtida a segmentação final da imagem.

O MRF criado nesta etapa promove a modelagem da distribuição dos rótulos das classes, sendo assim, o objetivo final da segmentação é a obtenção de uma amostra desse MRF que se aproxime da segmentação ótima. Tal amostra é obtida por meio do método de relaxação ICM, o qual minimiza a energia do sistema de maneira determinística. Visando à redução do impacto causado pela convergência para regiões de mínimos locais, um passo adicional, denominado *perturbação*, é acrescentado no ICM.

Função de Energia

Como descrito anteriormente, neste trabalho a segmentação baseada em dependência espacial é efetuada com a utilização do algoritmo ICM, para isso, alguns parâmetros devem ser obtidos antes do início de sua execução. Tais parâmetros, entre eles os descritores para cada pixel, a distribuição de características e o MRF, são utilizados para definir a função de energia.

Na seção 2.3 foi descrito que o algoritmo ICM visa à minimização da probabilidade *a posteriori* mostrada na equação 3.16, para isso, devem ser obtidos os componentes $P(Y_s = y_s | X_s = x_s)$ e $P(X_s = x_s | \partial_s)$, que representam, respectivamente, a distribuição de características dos pixels e a probabilidade *a priori*. Tais componentes serão derivados a

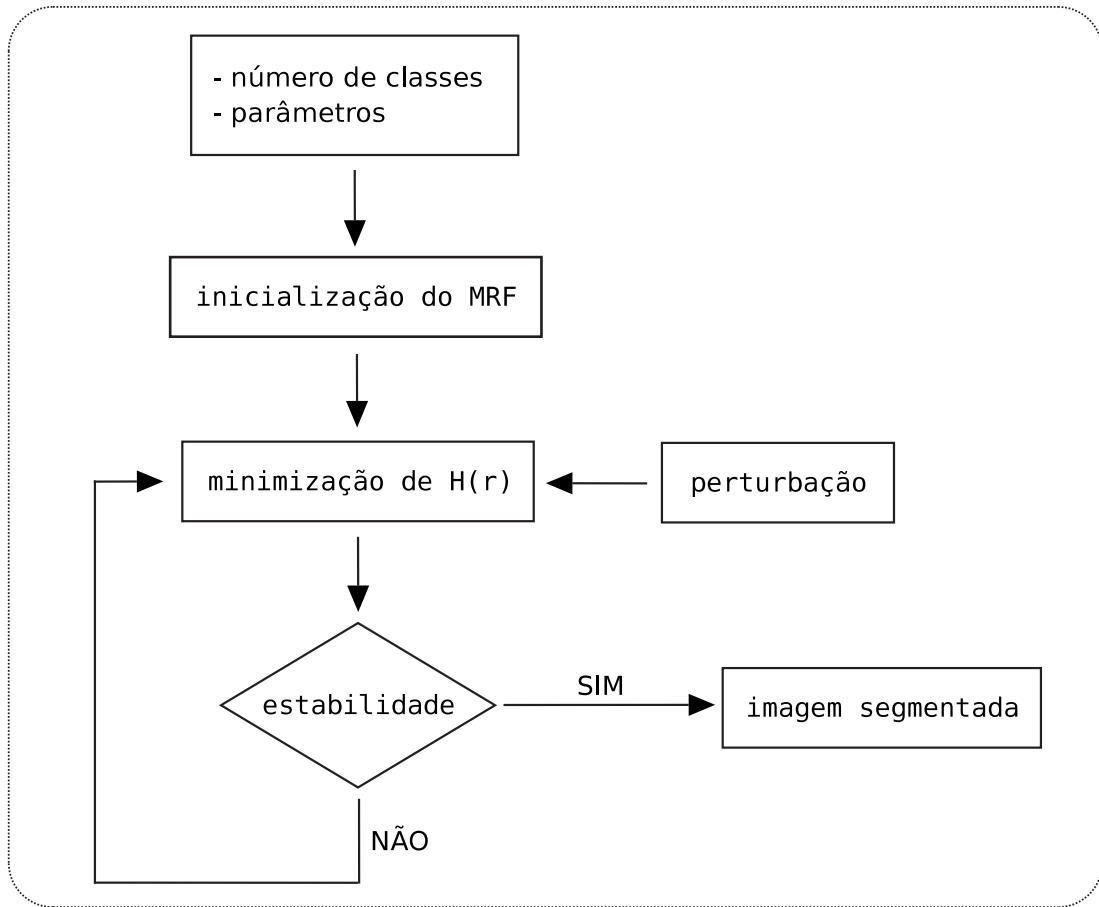


Figura 3.6: Diagrama ilustrando a segmentação baseada em pixels. A segmentação final é obtida a partir da realização do MRF que modela a distribuição dos rótulos das classes.

seguir, porém de forma indireta, pois a maximização de $P(X_s = x_s | Y_s = y_s)$ será transformada na minimização de uma função de energia local.

$$P(X_s = x_s | y, \partial_s) \propto P(Y_s = y_s | X_s = x_s) P(X_s = x_s | \partial_s) \quad (3.16)$$

Utilizando o resultado descrito na equação 2.95 sobre a distribuição da probabilidade *a posteriori* de um MRF e baseando-se no fato que a estimação MAP tem objetivo de maximizar $P(\mathbf{X}|\mathbf{Y})$ encontrando uma realização \mathbf{x} , essa maximização pode ser vista como a minimização da função de energia $H(\mathbf{x}|\mathbf{y})$. Dado que a função de partição Z representa um termo em comum para todas as realizações, quanto maior o valor de $H(\mathbf{x}|\mathbf{y})$ menor será a probabilidade *a posteriori*, repetida aqui na equação 3.17 para facilitar a explicação subsequente.

$$P(\mathbf{X}|\mathbf{Y}) = \frac{1}{Z} \exp(-H(\mathbf{x}|\mathbf{y})) \quad (3.17)$$

Com objetivo de utilizar apenas a função de energia, vai-se desenvolver o lado direito do teorema de Bayes (equação 2.94), sem o termo em comum $P(\mathbf{y})$, de modo a encontrar uma função de energia global que possa substituir $H(\mathbf{x}|\mathbf{y})$ na equação 3.17. Considerando a condição de independência entre as variáveis aleatórias que compõem a distribuição de características, imposta na seção 2.3, e tendo que a distribuição *a priori* de um MRF é dada pela equação 2.10, obtém-se a equação 3.18 como sendo a probabilidade *a posteriori*, onde n denota o número de variáveis aleatórias presentes no MRF (o mesmo que o número de pixels contidos na imagem).

$$\begin{aligned}
 P(\mathbf{X}|\mathbf{Y}) &= \frac{1}{Z} \exp(-H_U(\mathbf{x})) \prod_{i=1}^n P(Y_i = \mathbf{y}_i | X_i = \mathbf{x}_i) \\
 &= \frac{1}{Z} \exp(-H_U(\mathbf{x})) \exp \left[\ln \left\{ \prod_{i=1}^n P(Y_i = \mathbf{y}_i | X_i = \mathbf{x}_i) \right\} \right] \\
 &= \frac{1}{Z} \exp(-H_U(\mathbf{x})) \exp \left[\sum_{i=1}^n \ln \{P(Y_i = \mathbf{y}_i | X_i = \mathbf{x}_i)\} \right] \\
 &= \frac{1}{Z} \exp \left(-H_U(\mathbf{x}) + \sum_{i=1}^n \ln \{P(Y_i = \mathbf{y}_i | X_i = \mathbf{x}_i)\} \right) \quad (3.18)
 \end{aligned}$$

Comparando a equação 3.18 com a equação 3.17, a função de energia a ser minimizada é mostrada na equação 3.19, onde $P(Y_i = \mathbf{y}_i | X_i = \mathbf{x}_i)$ denota a distribuição de probabilidade das características para i -ésima variável aleatória.

$$H(\mathbf{x}|\mathbf{y}) = H_U(\mathbf{x}) - \sum_{i=1}^n \ln \{P(Y_i = \mathbf{y}_i | X_i = \mathbf{x}_i)\} \quad (3.19)$$

Percebe-se que o termo $P(Y_i = \mathbf{y}_i | X_i = \mathbf{x}_i)$ não especifica uma distribuição de probabilidade. Portanto, conforme normalmente efetuado, considera-se que esse termo segue a distribuição Gaussiana multivariada, dessa maneira, pode-se continuar o desenvolvimento da equação anterior. Para maior clareza, esse termo é tratado isoladamente, obtendo-se a equação 3.20 como resultado, onde l representa o número de medidas que compõem o vetor de características e Σ_i e μ_i denotam, respectivamente, a matriz de covariância e o vetor de médias da i -ésima classe.

$$P(Y_i = \mathbf{y}_i | X_i = \mathbf{x}_i) = \frac{1}{(2\pi)^{l/2} |\Sigma_{\mathbf{x}_i}|^{1/2}} \exp \left[-\frac{1}{2} \{(\mathbf{y}_i - \mu_{\mathbf{x}_i})^T \Sigma_{\mathbf{x}_i}^{-1} (\mathbf{y}_i - \mu_{\mathbf{x}_i})\} \right]$$

$$\begin{aligned}
\ln\{P(Y_i = y_i | X_i = x_i)\} &= \ln\left\{\frac{1}{(2\pi)^{1/2}|\Sigma_{x_i}|^{1/2}}\right\} - \frac{1}{2}\{(y_i - \mu_{x_i})^T \Sigma_{x_i}^{-1}(y_i - \mu_{x_i})\} \\
&= -\ln\{(2\pi)^{1/2}|\Sigma_{x_i}|^{1/2}\} - \frac{1}{2}\{(y_i - \mu_{x_i})^T \Sigma_{x_i}^{-1}(y_i - \mu_{x_i})\} \\
&= -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(|\Sigma_{x_i}|) - \frac{1}{2}\{(y_i - \mu_{x_i})^T \Sigma_{x_i}^{-1}(y_i - \mu_{x_i})\} \quad (3.20)
\end{aligned}$$

Portanto, a função de energia global passa a apresentar a forma mostrada na equação 3.21.

$$H(\mathbf{x}|\mathbf{y}) = 2H_U(\mathbf{x}) + \sum_{i=1}^n \{\ln(2\pi) + \ln(|\Sigma_{x_i}|) + (y_i - \mu_{x_i})^T \Sigma_{x_i}^{-1}(y_i - \mu_{x_i})\} \quad (3.21)$$

Ainda falta definir a função de potencial do MRF. Considerando cliques que apresentam apenas dois componentes, tem-se a função potencial mostrada na equação 3.22, denominada *M-level* MRF (Besag 1986), onde C_r denota o conjunto formado pelos cliques que contêm a variável aleatória r e β representa a intensidade do relacionamento entre os vizinhos, neste trabalho, determinado experimentalmente.

$$U(x_r, x_s) = \begin{cases} \beta, & \text{se } x_r \neq x_s \quad \forall x_s \in C_r \\ 0, & \text{caso contrário} \end{cases} \quad (3.22)$$

Dessa maneira, a função de potencial para todas as variáveis do MRF apresenta a forma mostrada na equação 3.23, onde $\delta(t) = 1$ se $t = 0$ e $\delta(t) = 0$ caso contrário, para $r = 1, 2, \dots, n$.

$$U(\mathbf{x}) = \beta \sum_{s \in C_r} \{1 - \delta(x_r - x_s)\} \quad (3.23)$$

Finalmente, a equação 3.24 apresenta a função de energia global. Tal resultado é conhecido na literatura (Jackson e Landgrebe 2002), no entanto, foi aqui apresentado com o objetivo de expressar a possibilidade que outras distribuições de características ou outros modelos de MRF sejam utilizados e, principalmente, restrições dependentes da aplicação possam ser impostas diretamente na função que, do mesmo modo o objetivo da segmentação continuaria sendo a minimização da função de energia.

$$H(\mathbf{x}|\mathbf{y}) = 2\beta \sum_{s \in C_r} \{1 - \delta(x_r - x_s)\} + \sum_{i=1}^n \{\ln(2\pi) + \ln(|\Sigma_{x_i}|) + (y_i - \mu_{x_i})^T \Sigma_{x_i}^{-1}(y_i - \mu_{x_i})\} \quad (3.24)$$

Pelo teorema 2.1 um MRF é um NGRF, desta maneira, dadas todas as variáveis aleatórias do MRF, a energia de cada uma delas depende apenas de sua vizinhança, pois $U_A = 0$ se A não for um clique, onde A denota um subconjunto das variáveis aleatórias. Considerando-se também a independência entre as variáveis que compõem a distribuição de características, a energia local de uma variável pode ser descrita pela equação 3.25. Nessa equação, k representa o número de vizinhos de r que estão atribuídos a classes distintas deste, o que vem diretamente da função de potencial mostrada na equação 3.22.

$$H(r) = 2k\beta + \ln(2\pi) + \ln(|\Sigma_{x_r}|) + (\mathbf{y}_r - \mu_{x_r})^T \Sigma_{x_r}^{-1} (\mathbf{y}_r - \mu_{x_r}) \quad (3.25)$$

Dessa maneira, o algoritmo ICM será utilizado para minimizar a função de energia local mostrada na equação 3.26, onde $s = 1, 2, \dots, n$, o que equivale a maximização da probabilidade local mostrada na equação 3.16 (Winkler 2003).

$$x_s = \min_{1 \leq r \leq m} \{2k\beta + \ln(|\Sigma_r|) + (\mathbf{y}_s - \mu_r)^T \Sigma_r^{-1} (\mathbf{y}_s - \mu_r)\} \quad (3.26)$$

Percebe-se que a equação 3.26 necessita das características de cada variável aleatória, dessa maneira, ainda devem ser calculados os descritores das características dos pixels. Para isso, a partir de janelas apresentando o mesmo tamanho daquelas empregadas para amostragem das regiões contendo classes homogêneas, são extraídas medidas acrescentadas nos vetores de características utilizados para a descrição de cada pixel. Tais vetores são denotados como o parâmetro \mathbf{y}_s da equação 3.26.

A extração dos descritores para cada pixel é a etapa que demanda maior tempo computacional durante todo o processo de segmentação, portanto, devem ser utilizados métodos eficientes de análise de texturas.

Inicialização do MRF

Após a definição da função $H(r)$, mostrada na equação 3.25, o ICM deve ser executado para que uma amostra do MRF que modela a distribuição dos rótulos seja obtida. Entretanto, antes de sua inicialização é necessário que seja estabelecida uma realização inicial para o MRF. Neste trabalho são descritas três abordagens: inicialização aleatória, resultado do classificador quadrático e regiões contendo classes homogêneas.

A inicialização aleatória é estabelecida com a simples atribuição aleatória das variáveis aleatórias do MRF, cada uma das variáveis recebe valores entre 1 e m , onde m denota número de classes. Geralmente, os trabalhos evitam essa forma de inicialização, pois leva

mais facilmente a regiões de mínimos locais. Entretanto, quando se utiliza a perturbação esse problema é reduzido (alguns resultados são apresentados na seção 4.2).

Uma vez estimados os parâmetros para descrição de cada classe, pode-se utilizar o classificador quadrático aplicado a todos os pixels para se obter uma amostra inicial do MRF. Sem considerar a probabilidade *a priori*, esse classificador é obtido apenas com a eliminação do termo $2k\beta$ da equação 3.26. Entretanto, se considerada a probabilidade *a priori*, tal classificador apresenta a forma mostrada na equação 3.27, resultado da combinação do teorema de Bayes com a distribuição Gaussiana multivariada.

$$x_s = \max_{1 \leq r \leq m} \left\{ \ln P_r - \frac{1}{2} \ln (|\Sigma_{x_r}|) - \frac{1}{2} ((y_s - \mu_r)^T \Sigma_r^{-1} (y_s - \mu_r)) \right\} \quad (3.27)$$

A última abordagem para inicialização do MRF utiliza as regiões contendo classes homogêneas em conjunto com uma das duas abordagens descritas anteriormente, essas aplicadas para efetuar a inicialização das variáveis aleatórias localizadas no restante da imagem. Embora o custo computacional para o classificador quadrático não seja elevado, pois todos os parâmetros necessários encontram-se calculados, a realização inicial obtida pela associação das regiões contendo classes homogêneas com a inicialização aleatória é aquela que apresenta o menor custo computacional. A figura 3.7 mostra a realização inicial para um MRF.

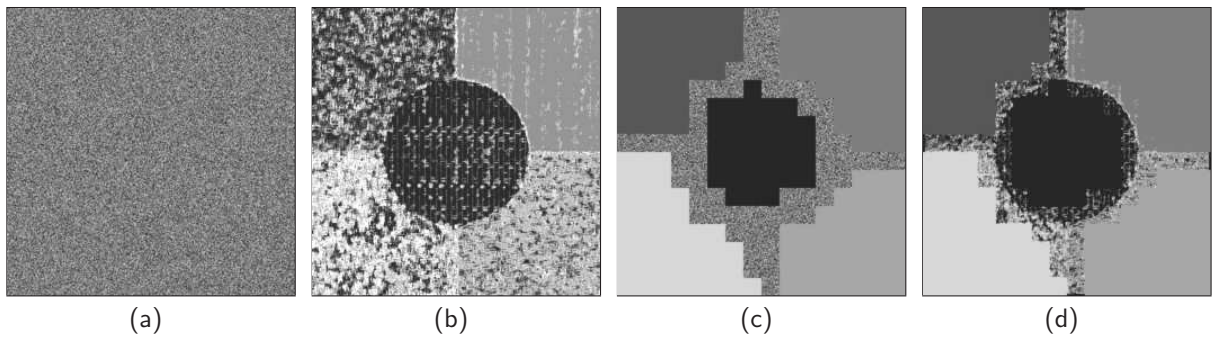


Figura 3.7: Realização inicial do MRF para a imagem mostrada na figura 3.4. (a) inicialização aleatória; (b) inicialização com o classificador quadrático, desconsiderando a probabilidade *a priori*; (c) regiões contendo classes homogêneas com inicialização aleatória; (d) regiões contendo classes homogêneas com o classificador quadrático, desconsiderando a probabilidade *a priori*.

Execução do ICM

Para que a segmentação final seja obtida o ICM deve ser executado proporcionando a minimização da função de energia $H(r)$. Para isso, ainda deve ser definida a perturbação, estipulada uma condição de parada e tratado da atualização da probabilidade *a priori* a cada ciclo do ICM.

Devido às alterações que ocorrem nas variáveis do MRF durante o ICM, torna-se necessário que a probabilidade *a priori*, neste caso substituída pelo número de vizinhos atribuídos a classes distintas, representado pelo parâmetro k , seja recalculada. Para isso, dada uma variável aleatória de interesse, deve-se determinar quantas de suas variáveis vizinhas estão atribuídas às classes distintas desta variável, atribuindo o resultado ao parâmetro k da equação 3.26.

Conforme comentado anteriormente, o ICM normalmente converge para mínimos locais, já que se baseia apenas na minimização local da energia e não permite que estados de maior energia sejam atingidos. Com o objetivo de reduzir o impacto causado pela convergência para regiões de mínimos locais, é proposta a inclusão de um novo passo no ICM, denominado *perturbação*, o qual possibilita a obtenção de estados que apresentem maior energia.

A idéia do acréscimo da perturbação no ICM, assim como em alguns trabalhos descritos na seção 2.4 que tentam ponderar a contribuição da distribuição dos rótulos e das características, é permitir que a minimização da energia seja mais dependente dos dados, nesse caso, ainda em maior grau por depender diretamente de como estão dispostas as fronteiras entre as classes.

Esse passo é acrescentado ao final de cada ciclo do ICM, ou seja, após a minimização da função de energia para todas as variáveis aleatórias. O algoritmo 3.3 apresenta uma descrição da perturbação que, de forma geral, atribui rótulos de maneira aleatória para variáveis contidas em regiões que apresentam classes isoladas.

Perturbação do ICM

Passo acrescentado ao final do algoritmo ICM. A função $c(x, y)$ retorna a classe a qual pertence o pixel localizado na coordenada (x, y) da imagem, sendo esta composta de $n = N \times N$ pixels. A função $\text{random}(m)$ retorna uma amostra entre 1 e m da distribuição de probabilidade uniforme. Com o acréscimo deste passo, quando uma região isolada apresentar altura ou largura menor que d , classes escolhidas de maneira aleatória são atribuídas aos pixels nela localizados. A variável d assume valores maiores conforme o avanço da execução, evitando que regiões isoladas apresentando áreas mais expressivas sejam ignoradas pela perturbação.

```

1  // testa se deve efetuar a perturbação no ciclo atual
2  se mod (ciclo, 10) = 0 então
3
4      // determina a largura ou altura da região isolada
5       $d \leftarrow 5 + \log(1 + \text{ciclo})$ 
6
7      // percorre toda a imagem
8      para  $x = 0$  até  $N - 1$  faça
9          para  $y = 0$  até  $N - 1$  faça
10
```

```

11      // regiões isoladas na horizontal
12      se  $c(x, y + d) \neq c(x, y - d)$  e  $c(x, y) \neq c(x, y + d)$  e
13          $c(x, y) \neq c(x, y - d)$  então
14          $c(x, y) \leftarrow \text{random}(m)$ 
15
16      // regiões isoladas na vertical
17      se  $c(x + d, y) \neq c(x - d, y)$  e  $c(x, y) \neq c(x + d, y)$  e
18          $c(x, y) \neq c(x - d, y)$  então
19          $c(x, y) \leftarrow \text{random}(m)$ 

```

Algoritmo 3.3: Perturbação do ICM.

A possibilidade de atingir estados de maior energia com a perturbação, assim como os algoritmos de relaxação combinados com SA, reduz a influência que as condições iniciais exercem na segmentação final, no entanto, nesse caso sem apresentar alto custo computacional como quando se utiliza SA. A seção 4.2 apresenta algumas discussões sobre o acréscimo da perturbação no ICM, inclusive mostrando um gráfico para o decaimento da energia do sistema.

Como condição de parada para o ICM, pode-se utilizar o número de ciclos, como normalmente efetuado, ou uma condição baseada na minimização da função de energia. Nessa última, quando não há variação significativa na função de energia entre dois ciclos consecutivos, considera-se que a segmentação final foi obtida. Após o final da execução do ICM, utiliza-se a realização do MRF para determinar a localização das fronteiras existentes entre as classes, como ilustrado na figura 3.8.

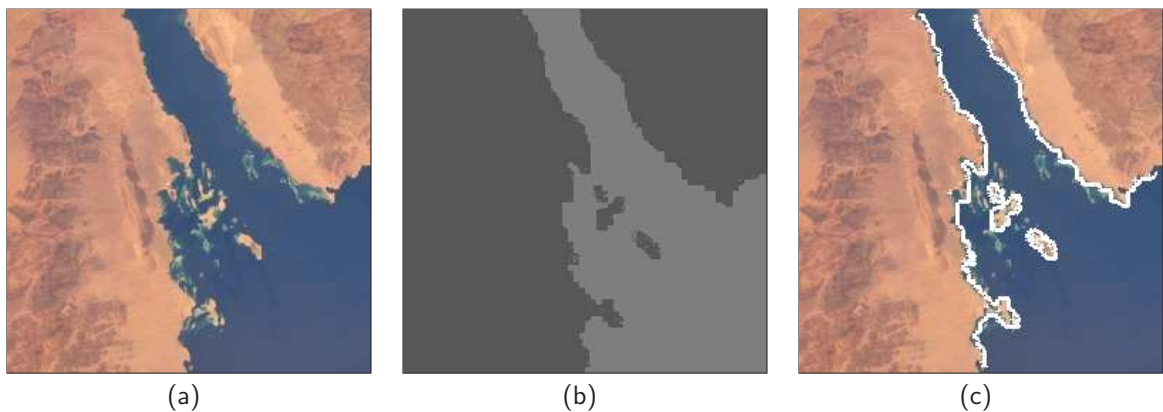


Figura 3.8: Segmentação final. (a) imagem original; (b) MRF obtido após o ICM; (c) fronteiras de classes indicadas sobre a imagem original.

CAPÍTULO 4

RESULTADOS EXPERIMENTAIS

Este capítulo apresenta uma série de experimentos visando à avaliação da metodologia proposta. Inicialmente, na seção 4.1, efetua-se a seleção dos métodos de análise de texturas descritos na seção 2.2, para tal, aplica-se a classificação de imagens em amostras de texturas. O método de análise de texturas que apresentar os resultados mais satisfatórios durante a classificação será utilizado na seção 4.2, que trata da avaliação do método de segmentação proposto.

Em cada uma das seções, apresenta-se inicialmente uma série de resultados por meio de gráficos ou imagens obtidas a partir da segmentação, com intuito de delinear a análise dos aspectos mais relevantes envolvidos nos processos de extração de características e segmentação.

4.1 Seleção das Características de Texturas

Para selecionar os métodos apresentados na seção 2.2 efetua-se uma série de experimentos baseados na classificação de texturas. Embora existam alguns trabalhos que efetuem comparações entre as características de texturas, tais como os propostos por Connors e Harlow (1980), Ohanian e Dubes (1992), Ojala et al. (1996), Randen e Husoy (1999), Sharma (2001), Singh e Singh (2002), Weszka et al. (1976) e Yong (2002), nenhum compara especificamente os métodos descritos na referida seção.

Esses experimentos visam determinar quais métodos são mais propícios de serem aplicados na etapa de segmentação, considerando os aspectos de acuracidade e tempo de execução. Sem observar este compromisso, a aplicação do método de segmentação pode ser inviabilizada por apresentar resultados impróprios ou alto custo computacional.

Assim como na maioria dos trabalhos, a abordagem proposta para seleção desses métodos se baseia na classificação de amostras de texturas. Devido a sua aplicação durante a minimização da função de energia na segmentação, embora de maneira indireta, este trabalho utiliza o classificador quadrático. Sua formulação é apresentada na equação 3.27, entretanto, ao invés de representar apenas as características de um pixel, como na seção 3.3, o termo y_s descreve uma amostra de textura.

Os resultados apresentados nos gráficos são obtidos a partir do valor médio calculado para

os conjuntos de imagens mostrados nas figuras 4.1 e 4.2, compostos de texturas pertencentes aos álbuns Brodatz (Brodatz 1966) e MeasTex (MeasTex 2005), respectivamente. Com o objetivo de obter valores estáveis, para cada experimento o método de classificação é executado cinco vezes e utiliza-se o valor médio dos resultados obtidos.

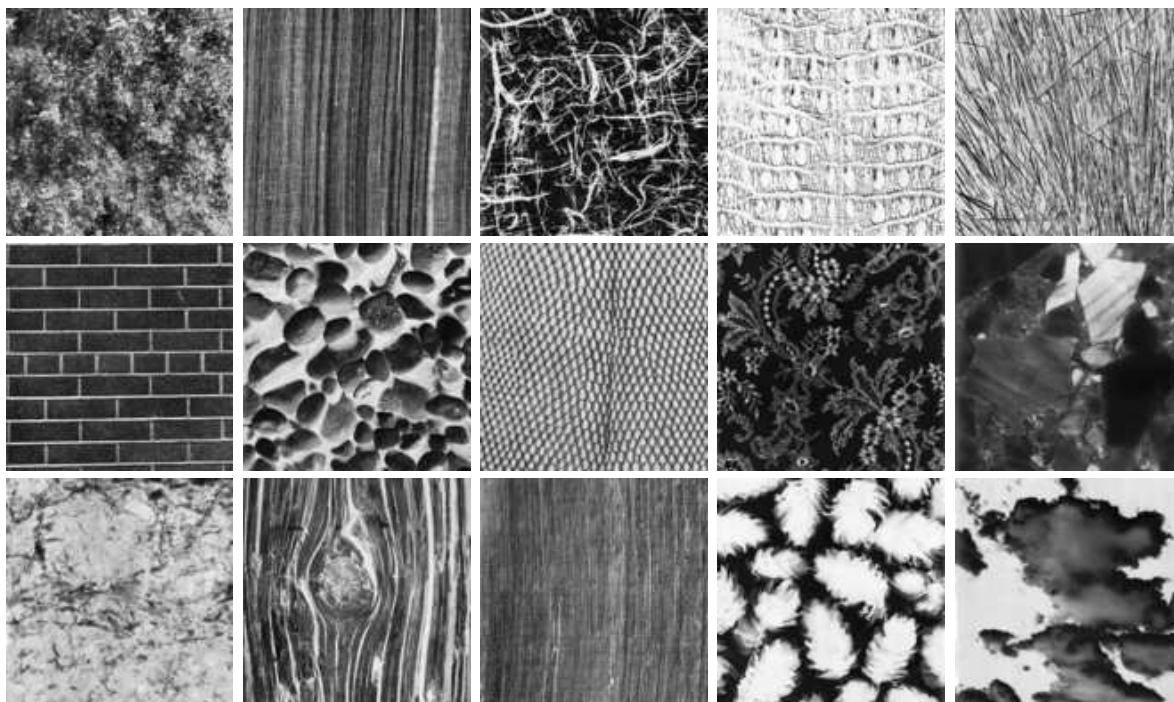


Figura 4.1: Imagens utilizadas para avaliação das características de texturas. Conjunto *Brodatz*, composto pelas imagens D100, D105, D107, D10, D15, D26, D31, D35, D40, D59, D63, D72, D79, D88 e D91.

Nesses conjuntos, cada textura é composta por 512×512 pixels e 256 tons de cinza, exceto no experimento cujos resultados são mostrados na figura 4.8(b) onde considera-se variações no número de tons de cinza das texturas. A partir de cada uma são obtidas 150 amostras escolhidas aleatoriamente, dentre as quais 100 são utilizadas para obtenção do conjunto de treinamento e 50 são posteriormente classificadas.

Mesmo com o processo aleatório para obtenção das texturas que compõem cada um dos conjuntos, deve-se salientar que os resultados aqui apresentados podem ser distintos daqueles obtidos a partir de dados provenientes de áreas específicas de aplicação. Tal dificuldade é encontrada em outros estudos comparativos, e seria superada apenas se comparações teóricas fossem efetuadas, entretanto, tais comparações são inviabilizadas devido à inexistência de uma definição de textura que seja amplamente aceita (Smith e Burns 1997).

A seguir são descritos os detalhes de implementação para os métodos de análise de texturas, a descrição de cada gráfico e, finalmente, os resultados obtidos são analisados.

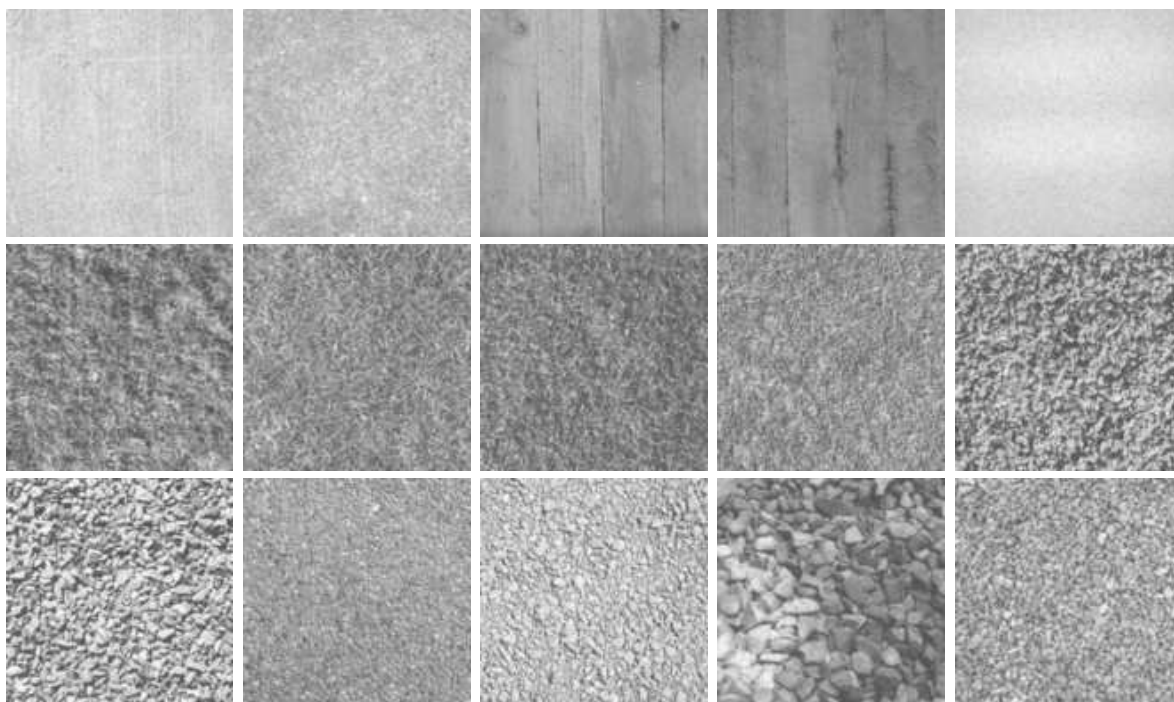


Figura 4.2: Imagens utilizadas para avaliação das características de texturas. Conjunto *Meas-
Tex*, composto pelas imagens Concrete.0000, Concrete.0002, Concrete.0006, Concrete.0007, Con-
crete.0010, Grass.0000, Grass.0002, Grass.0008, Grass.0011, Grass.0016, Rock.0004, Rock.0007,
Rock.0011, Rock.0019 e Rock.0023.

Detalhes de Implementação

Todos os experimentos efetuados foram executados em microcomputador PC Pentium IV 1600 MHz com 256 Mbytes de memória RAM utilizando linguagem de programação C com o sistema operacional Linux. A seguir são descritos os detalhes de implementação para os métodos de análise de texturas, pois para alguns deles não foram consideradas todas as medidas apresentadas na seção 2.2.

Na implementação do modelo generalizado de Ising considera-se as vizinhanças de primeira, segunda, terceira e quarta ordens, tendo número de parâmetros igual a 3, 5, 7 e 9, respectivamente. A ordem da vizinhança que apresenta melhor resultado para classificação é avaliada experimentalmente.

Para o método de codificação de características de texturas são consideradas a aspereza, homogeneidade, convergência média, variância, entropia de código e a similaridade de código. Determina-se o valor de Δ a partir de experimentos, nos quais efetua-se a classificação utilizando valores distintos para o referido parâmetro.

A função de autocorrelação implementada possui dois parâmetros a serem ajustados, p e q , ambos diretamente relacionados ao número de medidas a serem acrescentadas no vetor de

características. Nos experimentos efetua-se variações nesses parâmetros, considerando que os valores de p e q são iguais.

Para o método baseado em unidade de textura são implementadas as medidas de simetria preto-branco, simetria geométrica e o grau de direção. Já para decomposição de valor singular implementa-se a média e o desvio padrão obtidos a partir dos valores singulares extraídos para cada amostra. Para determinar o tamanho da janela e o número de valores singulares a serem extraídos efetua-se dois experimentos.

A partir da matriz de co-ocorrência são extraídas as medidas de entropia, segundo momento angular, contraste, correlação, variância e homogeneidade, sendo que o vetor de características é composto pela média de cada medida para as orientações de 0° , 45° , 90° e 135° . Para determinar o parâmetro d são testados valores entre 1 e 6.

No método baseado em matrizes de comprimento de corridas de cinza são implementadas as medidas originalmente propostas por Galloway (1975), ênfase em corridas curtas, ênfase em corridas longas, não uniformidade de tom de cinza, não uniformidade de tamanho de corrida e percentagem de corrida. Tais medidas são extraídas a partir das matrizes determinadas para corridas apresentando orientações de 0° e 90° , resultando, portanto, em 10 características de texturas.

Para o método baseado em transformada wavelet, descrito na seção 2.2.2, considera-se a família de funções Daubechies 4, as quais efetuam uma decomposição em dois níveis, conforme a mostrada na figura 2.13(b). Os coeficientes de energia são obtidos a partir das sub-imagens que apresentam frequências mais altas, desta maneira, o vetor de características é composto de 6 medidas que resumem as propriedades das texturas analisadas.

Finalmente, os dois últimos métodos implementados baseiam-se em estatísticas de primeira ordem e no espectro de Fourier. Para o primeiro, calcula-se o desvio padrão, média, assimetria, curtose, energia e a entropia a partir do histograma de tons de cinza. Já no segundo, determina-se a variância e a média a partir da função $S(\theta)$, obtida diretamente do espectro de Fourier.

Experimentos Efetuados

Como descrito na seção 3.2, os tons de cinza que compõem cada janela amostrada podem ser utilizados diretamente para sua descrição, entretanto, além serem altamente correlacionados, o espaço de características apresenta dimensionalidade muito elevada, o que acarreta resultados impróprios e alto custo computacional. Com intuito de mostrar esses efeitos, os gráficos da figura 4.3 apresentam o desempenho obtido na classificação e o tempo requerido

por esta, onde o eixo das abscissas representa o tamanho da janela e o eixo das ordenadas denota, em (a), a classificação correta, normalizada entre 0 e 1, e o tempo mostrado em segundos em (b).

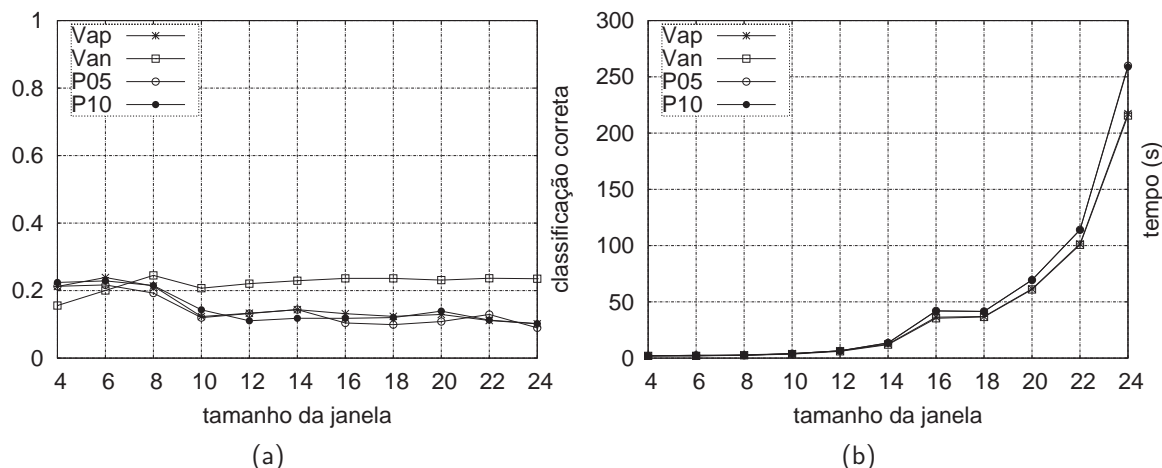


Figura 4.3: Classificação utilizando os próprios tons de cinza como característica de texturas, para variáveis padronizadas (Vap), variáveis sem padronização (Van), cinco primeiros componentes principais (P05) e para os dez primeiros componentes principais (P10). (a) classificação correta com variação no tamanho da janela; (b) tempo de execução conforme o tamanho da janela.

Alguns dos métodos descritos na seção 2.2 apresentam parâmetros que, conforme os valores assumidos proporcionam resultados distintos na classificação. Dentre esses métodos estão a estimação dos parâmetros pelo modelo generalizado de Ising, no qual a ordem da vizinhança influi (figura 4.4), alterações efetuadas no valor de Δ para o método de codificação de características de texturas (figura 4.5) e o tamanho das janelas utilizadas para os métodos de autocorrelação (figura 4.6) decomposição de valor singular (figura 4.7). Nesses experimentos, o eixo das abscissas e ordenadas indicam, respectivamente, o tamanho da janela e o índice de acerto na classificação obtida.

Para o método baseado em matrizes de co-ocorrência efetua-se alterações no parâmetro d e no número de tons de cinza que compõem a textura, conforme mostrado na figura 4.8. As alterações efetuadas no número de níveis de cinza estão diretamente relacionadas ao tempo requerido para a extração das características a partir da matriz de co-ocorrência, a qual apresenta tamanho proporcional ao referido número.

Com intuito de justificar a restrição imposta na seção 3.2, de que as janelas utilizadas para amostragem das regiões homogêneas e para extração de características dos pixels tenham o mesmo tamanho, apresenta-se dois gráficos, nos quais altera-se o tamanho das janelas para etapa de treinamento e classificação, indicados nos eixos x e y , respectivamente. Tal experimento é efetuado para o método baseado em matrizes de co-ocorrência (figura 4.9) e na

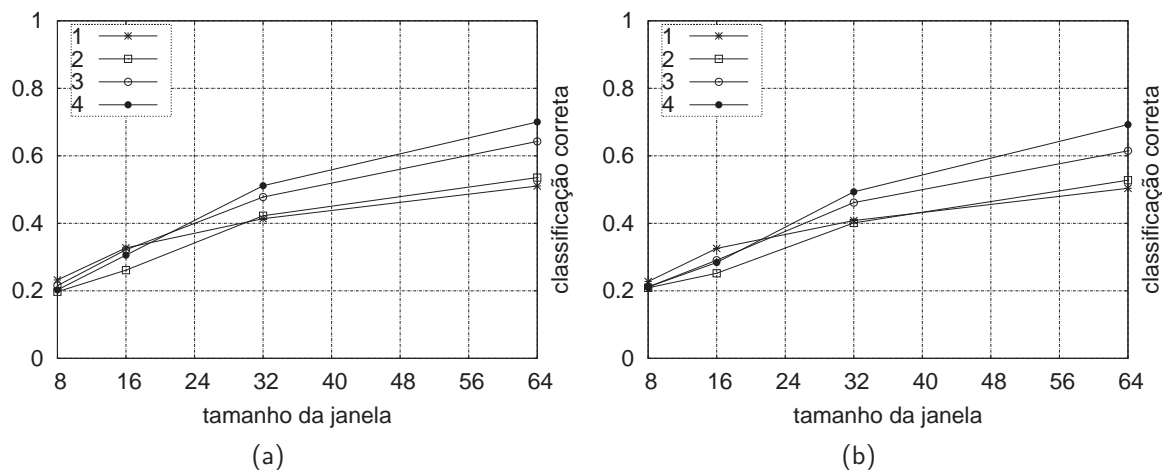


Figura 4.4: Classificação efetuada com alterações na ordem da vizinhança do método baseado no modelo generalizado de Ising. (a) variáveis padronizadas; (b) variáveis sem padronização.

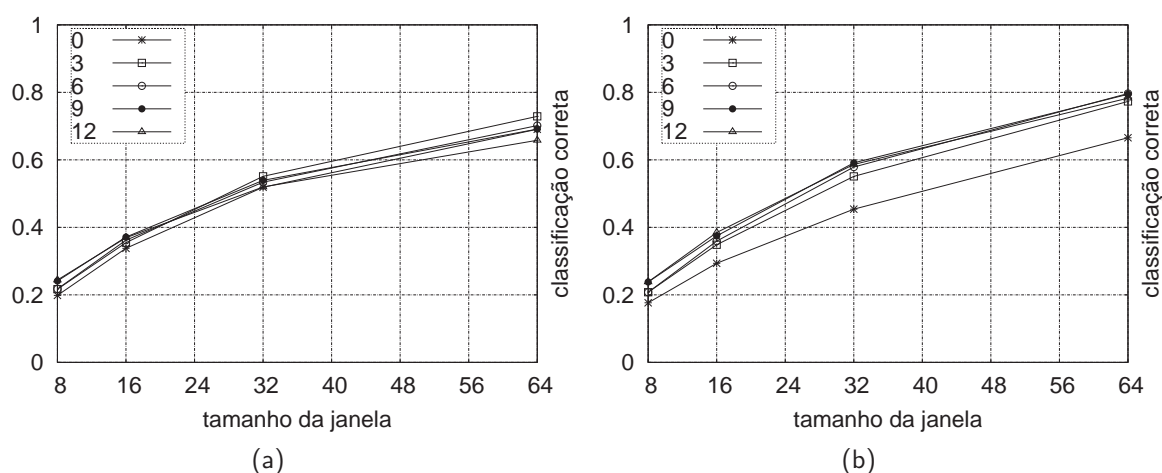


Figura 4.5: Classificação efetuada com alterações no parâmetro Δ do método de codificação de características de texturas. (a) variáveis padronizadas; (b) variáveis sem padronização.

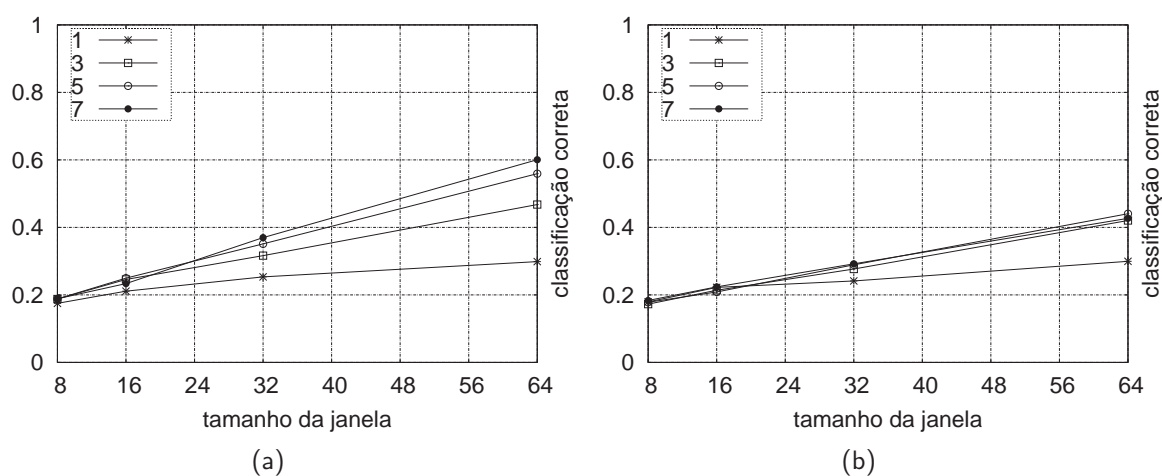


Figura 4.6: Classificação efetuada com alterações nos parâmetros p e q da função de autocorrelação. (a) variáveis padronizadas; (b) variáveis sem padronização.

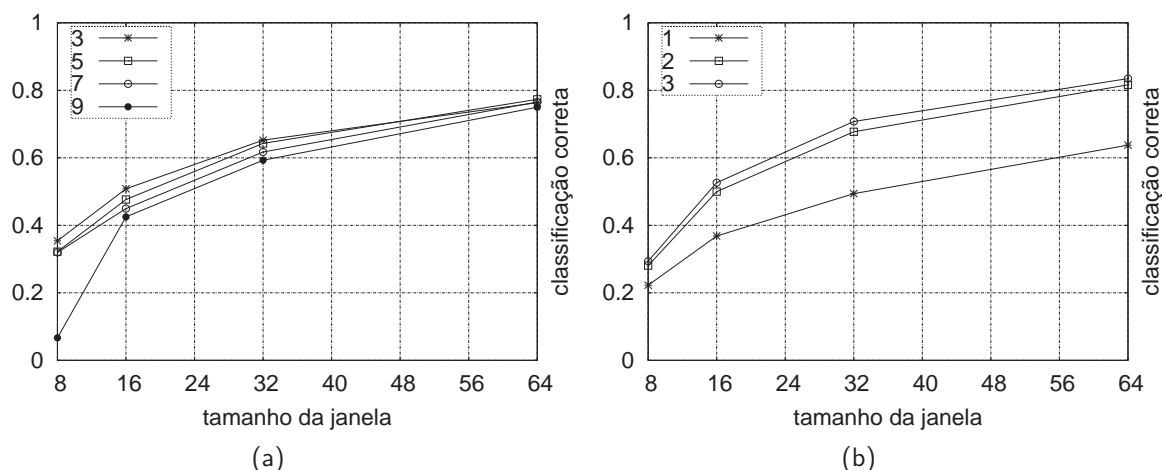


Figura 4.7: Classificação efetuada com alterações no parâmetros do método de decomposição de valor singular, utilizando-se variáveis padronizadas. (a) tamanho da janela utilizada para decomposição; (b) número de valores singulares utilizados.

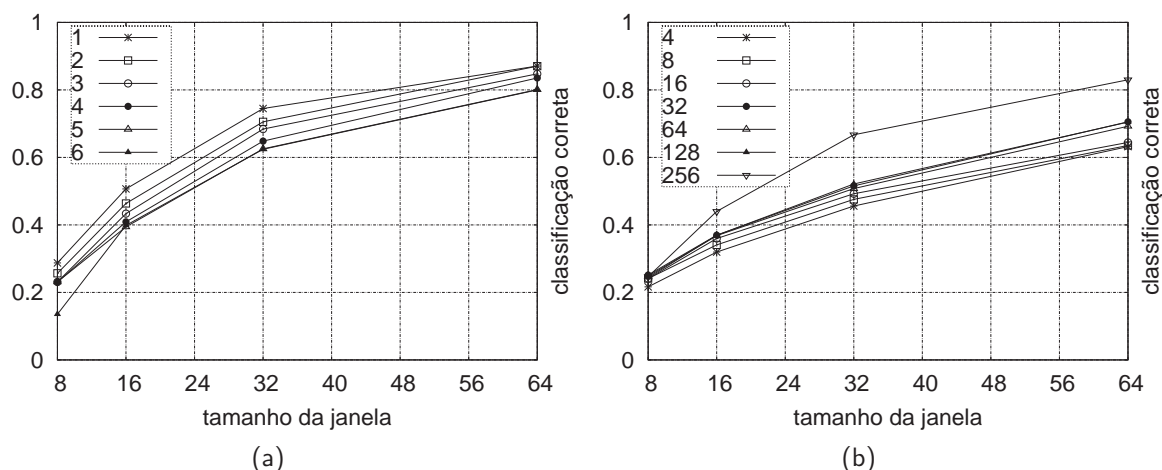


Figura 4.8: Classificação baseada em matrizes de co-ocorrência, utilizando-se variáveis padronizadas. (a) alteração no parâmetro d; (b) variação no número de tons de cinza presentes na imagem de entrada.

transformada wavelet (figura 4.10), onde o eixo z representa a classificação correta, indicada em tons de cinza na projeção mostrada no plano xy.

Utilizando-se os parâmetros que proporcionam o melhor desempenho nos experimentos cujos resultados são mostrados nas figuras 4.4 a 4.7, cria-se os gráficos mostrados nas figuras 4.11 e 4.12, onde compara-se o tempo requerido e o resultado obtido na classificação, respectivamente. Com intuito de facilitar a visualização, a comparação de tempo é dividida em dois grupos, o primeiro compreende os métodos mais rápidos e o segundo é composto por aqueles que apresentam maior custo computacional.

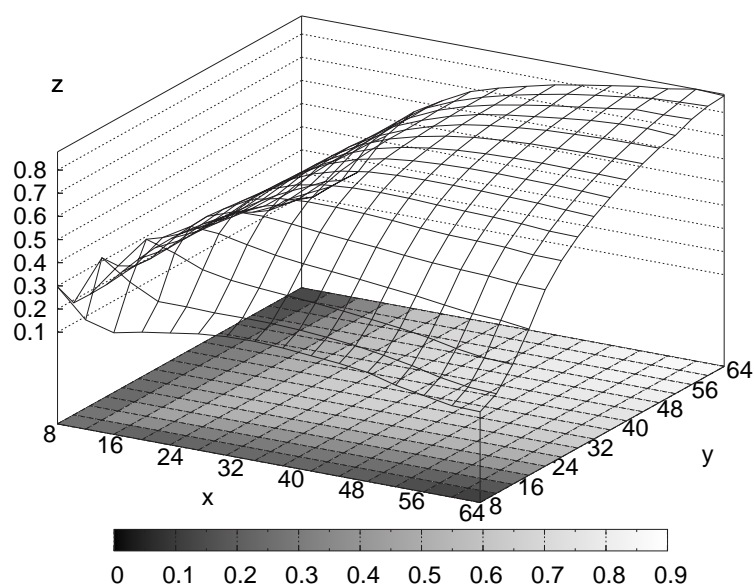


Figura 4.9: Classificação efetuada a partir de combinações distintas de tamanhos de janelas para classificação e treinamento, utilizando características extraídas a partir de matrizes de co-ocorrência.

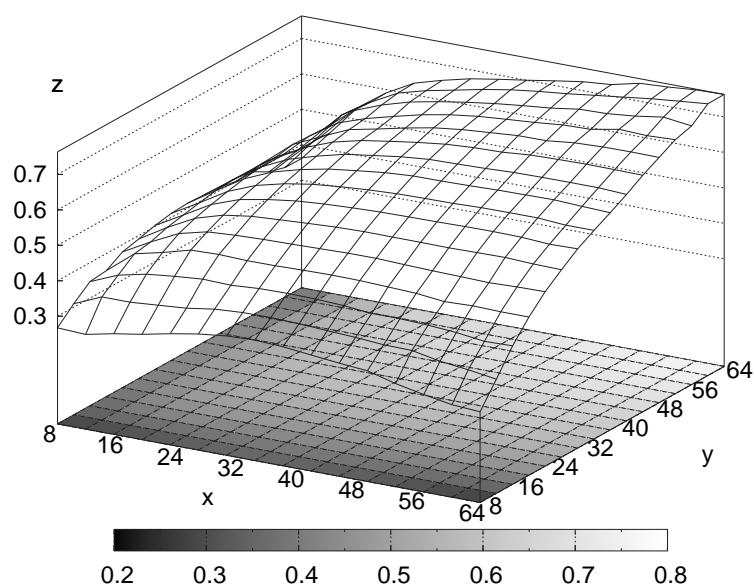


Figura 4.10: Classificação efetuada a partir de combinações distintas de tamanhos de janelas para classificação e treinamento, utilizando características obtidas a partir da transformada wavelet.

Análise dos Resultados Obtidos

Dentre os principais aspectos a serem analisados estão a utilização direta dos níveis de cinza; a alteração no valor dos parâmetros de alguns métodos de análise de texturas; a aplicação de janelas distintas nas etapas de treinamento e classificação; o comportamento do classificador

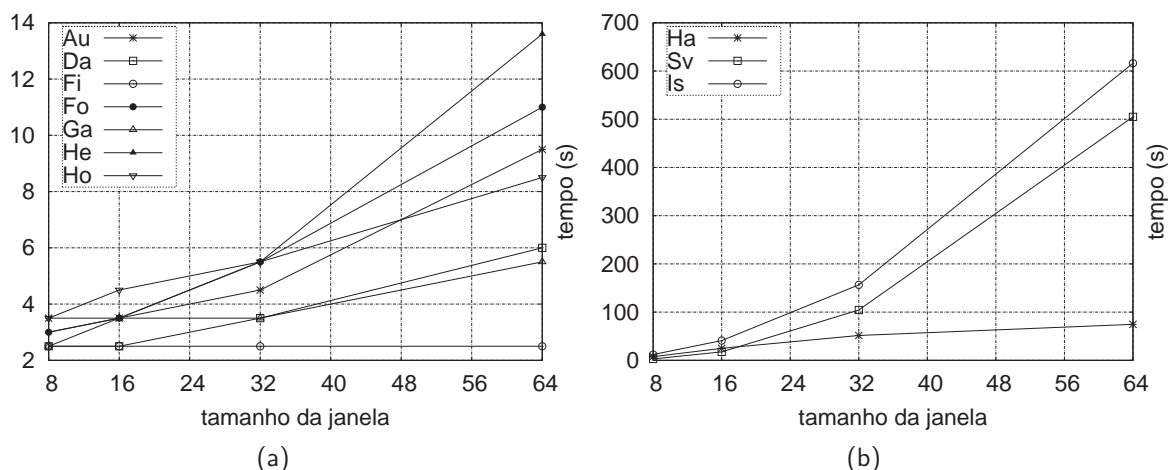


Figura 4.11: Tempo requerido para classificação, utilizando os métodos de autocorrelação (Au), transformada wavelet (Da), estatísticas de primeira ordem (Fi), espectro de Fourier (Fo), matrizes de comprimento de corridas de cinza (Ga), unidade de textura (He), codificação de características de texturas (Ho), matrizes de co-ocorrência (Ha), decomposição de valor singular (Sv) e modelo generalizado de Ising (Is). (a) métodos rápidos; (b) métodos lentos.

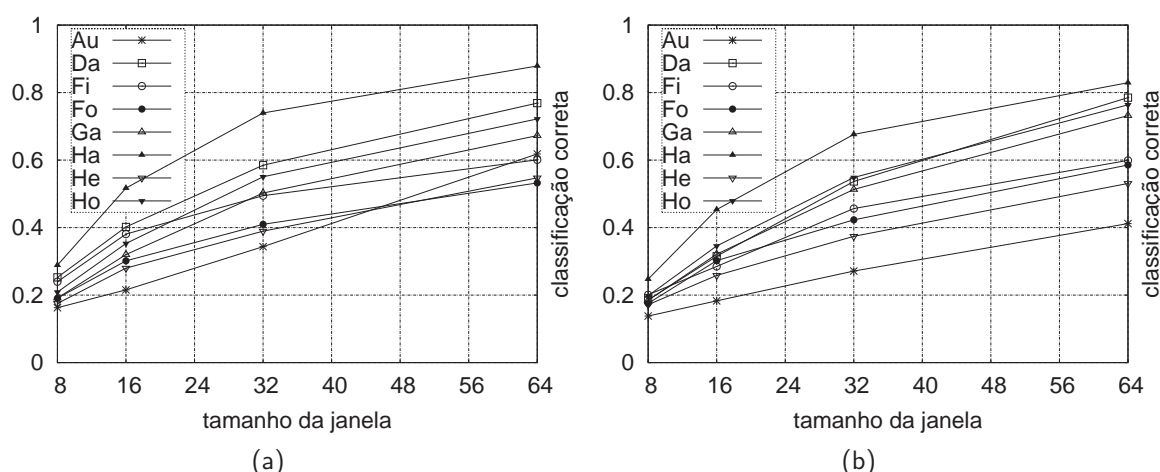


Figura 4.12: Classificação obtida (utiliza-se a mesma legenda da figura 4.11). (a) variáveis padronizadas; (b) variáveis sem padronização.

quadrático quando aplicado às variáveis sem padronização e, finalmente, a avaliação geral do desempenho obtido pelos métodos de análise de texturas considerados.

Além de ser uma abordagem utilizada pelo sistema visual humano, como descrito na seção 2.2, a aplicação de métodos para análise de texturas é justificada pelo fraco desempenho obtido a partir dos tons de cinza como características de texturas. A figura 4.3(a) mostra que a classificação correta alcançada com níveis de cinza é baixa, mesmo quando se utiliza variáveis padronizadas ou os primeiros componentes principais, neste caso ficando clara a necessidade de uma etapa de seleção de características, pois os resultados estão abaixo daqueles obtidos quando se utiliza um maior número de variáveis.

Dessa maneira, torna-se interessante a aplicação de métodos capazes de resumir as propriedades apresentadas pelas regiões da imagem e que possibilitem a redução na dimensionalidade dos dados, esta justificada diretamente pelo gráfico da figura 4.3(b) que, mesmo para janelas com tamanho pequeno, a aplicação direta dos tons de cinza apresenta alto custo computacional. Por exemplo, para janelas compostas de 8×8 pixels o espaço de características apresenta 64 dimensões, enquanto em métodos de análise de texturas o referido espaço normalmente não chega na metade deste valor, mesmo quando são descritas janelas maiores.

A escolha dos parâmetros que apresentam a melhor classificação utilizando-se o modelo generalizado de Ising, codificação de características de texturas, autocorrelação e decomposição de valor singular, é efetuada com base nos experimentos cujos resultados são mostrados nas figuras 4.4 a 4.7. Os parâmetros que, no geral, proporcionam os melhores resultados são utilizados para o experimento que compara o desempenho de todos os métodos de análise de texturas.

Para o modelo generalizado de Ising, no geral, os melhores resultados são obtidos quando se considera vizinhança de quarta ordem para o MRF. Embora a vizinhança de primeira ordem apresente bom desempenho quando estão sendo analisadas janelas com poucos pixels, como 8×8 e 16×16 , para janelas maiores, segundo os experimentos efetuados, quanto maior a ordem da vizinhança utilizadas melhores são os resultados obtidos.

Embora fique um pouco abaixo para janelas compostas de 16×16 pixels quando se utiliza variáveis padronizadas, a utilização da função de autocorrelação com os parâmetros p e q iguais a 7 proporciona os melhores resultados, acarretando em vetores de características que apresentam 49 medidas. Janelas maiores não foram avaliadas devido ao rápido aumento na dimensionalidade do espaço de características.

Utilizando-se o método de decomposição de valor singular, efetua-se experimentos para determinar o número de valores singulares e o tamanho da janela a partir da qual esses devem ser extraídos. Como resultados tem-se que 3 valores singulares obtidos a partir de janelas 3×3 e 5×5 proporcionam uma classificação acurada, equiparando-se, inclusive, com aquela obtida pelo método baseado em matrizes de co-ocorrência.

Os experimentos efetuados para determinar os melhores valores para o parâmetro Δ do método de codificação de características de texturas são os que apresentam maiores variações nos resultados obtidos. Quando se utiliza variáveis padronizadas para janelas com 32×32 e 64×64 pixels, $\Delta = 3$ proporciona os melhores resultados, entretanto, para todas as outras configurações $\Delta = 9$ é aquele que resulta na melhor classificação. Esse último valor é considerado durante a comparação entre os métodos.

Os gráficos da figura 4.8 mostram que a melhor classificação efetuada pelo método baseado em matrizes de co-ocorrência é obtida com $d = 1$ e quando se utiliza a imagem original, ou seja, composta por 256 tons de cinza. Dado que o custo computacional para essa configuração é alto, como mostra a figura 4.11(b), normalmente efetua-se uma transformação com intuito de reduzir o número de tons de cinza da imagem, com isso, embora ocorra perda na qualidade da classificação obtida, conforme mostrado na figura 4.8(b), o tempo requerido torna-se aceitável.

A partir de variações nos tamanhos das janelas utilizadas nas etapas de treinamento e classificação obtém-se os gráficos das figuras 4.9 e 4.10. Percebe-se que, em regiões onde x exibe os mesmos valores de y , a classificação apresenta melhores resultados, ou seja, quando se utiliza janelas com o mesmo tamanho durante o treinamento e classificação. Isto justifica a restrição imposta na seção 3.2, que indica a aplicação de janelas iguais para amostragem das regiões homogêneas e para extração de características dos pixels.

De maneira geral, a partir dos gráficos das figuras 4.4, 4.5 e 4.12, percebe-se que a classificação obtida tanto para variáveis padronizadas quanto sem padronização apresenta resultados semelhantes, o que tenderia a contradizer a justificativa apresentada na seção 3.2, de que os dados devem estar em uma mesma escala. Porém, neste caso, tais resultados são atribuídos à compensação efetuada pelo classificador quadrático, o qual considera a variância e a média das classes (Theodoridis e Koutroumbas 2003).

Como resultado final da avaliação das características de texturas apresenta-se os gráficos das figuras 4.11 e 4.12, os quais representam, respectivamente, a avaliação do tempo e da qualidade de classificação. Observando-se apenas a qualidade obtida na classificação, os métodos baseados em matriz de co-ocorrência, a decomposição de valor singular, mostrado na figura 4.7, e a transformada wavelet apresentam os melhores resultados, enquanto aqueles que se baseiam na função de autocorrelação, unidade de textura e espectro de Fourier obtêm os piores resultados.

Entretanto, pelo gráfico 4.11(b), percebe-se que os métodos baseados em matrizes de co-ocorrência e decomposição de valor singular necessitam de grande quantidade de computação. Desse modo, considerando o compromisso entre qualidade e tempo requerido, o método que apresenta melhor desempenho é o baseado em transformada wavelet, o qual será preferencialmente utilizado na obtenção dos resultados descritos na próxima seção.

Deve-se ressaltar o desempenho do método baseado em estatísticas de primeira ordem que, embora seja computado em tempo praticamente constante, como mostra a figura 4.11(a), compõe o grupo que apresenta resultados intermediários. Dessa maneira, encontra-se como uma boa alternativa para aplicações que priorizam o tempo de execução, porém, necessitam

que os resultados apresentem certa qualidade.

4.2 Avaliação da Metodologia Proposta

Esta seção apresenta os resultados obtidos com a aplicação da metodologia proposta no capítulo 3 para imagens coloridas e monocromáticas. Como método de análise de texturas, utiliza-se a extensão tridimensional da matriz de co-ocorrência descrita na seção 2.2.1 para imagens coloridas e o método baseado em transformada wavelet para segmentação de imagens monocromáticas, exceto para a imagem mostrada na figura 4.20.

Em todos os experimentos, considera-se $\beta = 1.5$, vizinhança de quarta ordem para o MRF, janelas para amostragem inicial compostas de 32×32 pixels para as imagens com tamanho 256×256 e 64×64 pixels para as imagens com tamanho 512×512 . Para amostragem das regiões homogêneas e na obtenção das características dos pixels, utiliza-se janelas compostas de 8×8 pixels. Já nas imagens da figuras 4.24 e 4.26, as quais apresentam tamanhos distintos dos anteriormente descritos, utiliza-se janelas compostas de 32×32 pixels para amostragem inicial.

Ao longo desta seção apresenta-se uma série de resultados, tanto aqueles obtidos pela aplicação da metodologia proposta quanto resultados auxiliares, tais como a comparação entre a segmentação baseada em dependência espacial com a segmentação tradicional, a detecção do número de classes por meio dos critérios apresentados na seção 3.2, a estabilidade da modificação do algoritmo *k-means* e o decaimento de energia proporcionado pela aplicação da perturbação.

Discussão dos Resultados

Dentre os aspectos relevantes a serem comentados estão a comparação efetuada entre os resultados obtidos quando não se considera a dependência espacial e quando se aplica o ICM (figuras 4.13 e 4.22); a aplicação da perturbação no ICM (figuras 4.14, 4.16, 4.18 e 4.20); a determinação do número de classes presentes na imagem (figuras 4.15, 4.17, 4.19, 4.21, 4.23 e 4.25); a estabilidade do método de agrupamento *k-means*, bem como sua instabilidade quando não se utiliza a padronização das variáveis (figura 4.27); o maior decaimento da energia quando se aplica a perturbação no ICM (figura 4.28) e algumas considerações finais sobre a metodologia proposta.

Como primeiro experimento compara-se o desempenho entre o método de segmentação proposto com e sem a utilização de dependência espacial, onde os resultados obtidos são

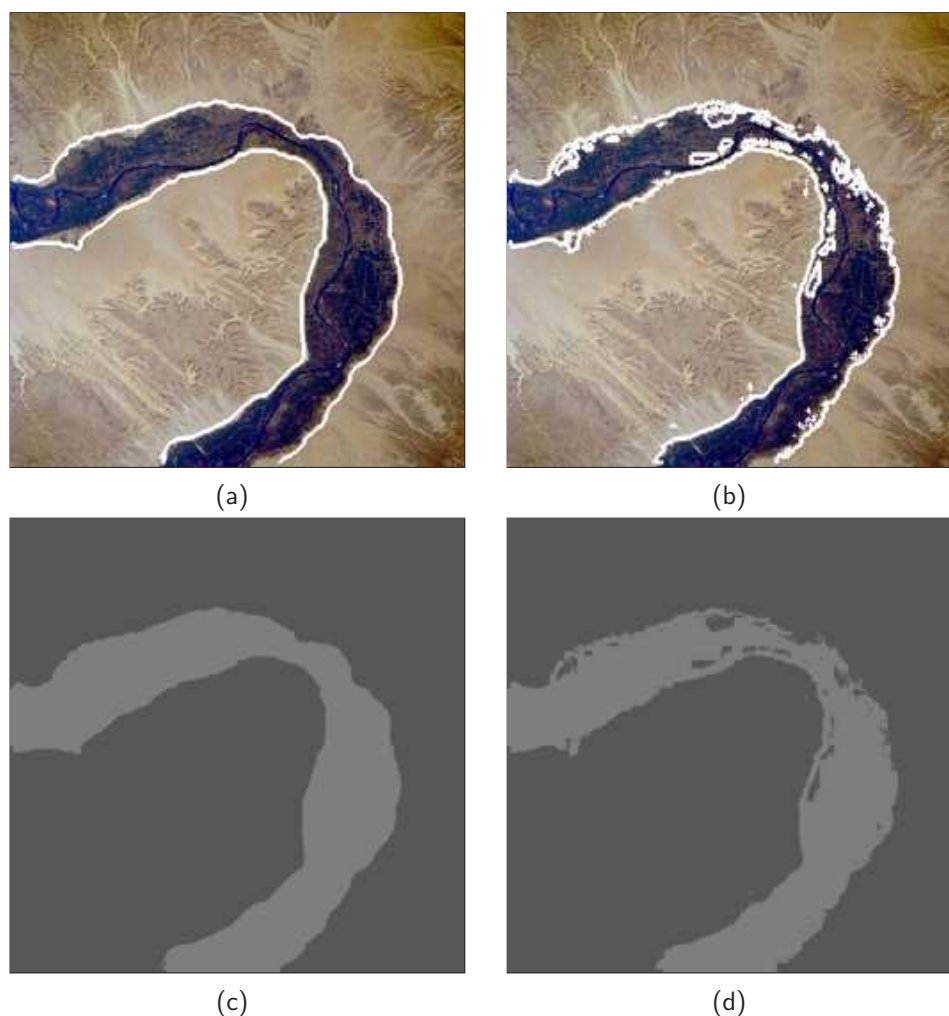


Figura 4.13: Segmentação resultante para imagem do rio Nilo, composta de 512×512 pixels. (a) resultado considerando a dependência espacial; (b) resultado utilizando apenas o classificador quadrático; (c) realização resultante do MRF, considerando a dependência espacial; (d) regiões resultantes da segmentação baseada no classificador quadrático.

mostrados nas figuras 4.13 e 4.22. Para que a dependência espacial seja ignorada, atribui-se o valor 0 ao parâmetro β da equação 3.26. Dessa maneira, a interação entre os vizinhos torna-se nula, obtendo-se, então, o classificador quadrático sem considerar a probabilidade *a priori*, conforme já comentado na seção 3.3.

A partir desses dois resultados, percebe-se que a consideração da dependência espacial proporciona maior adaptação aos contornos, principalmente na figura 4.13, que apresenta algumas reentrâncias. Outra característica observada está no fato de que o resultado obtido com o classificador quadrático, justamente por desconsiderar as classes dos elementos vizinhos, apresenta diversas regiões isoladas, além de considerar erroneamente algumas atribuições de classes, como na fronteira horizontal direita da figura 4.22.

Na seção 3.3 foi descrita a introdução da etapa de perturbação no algoritmo ICM com obje-

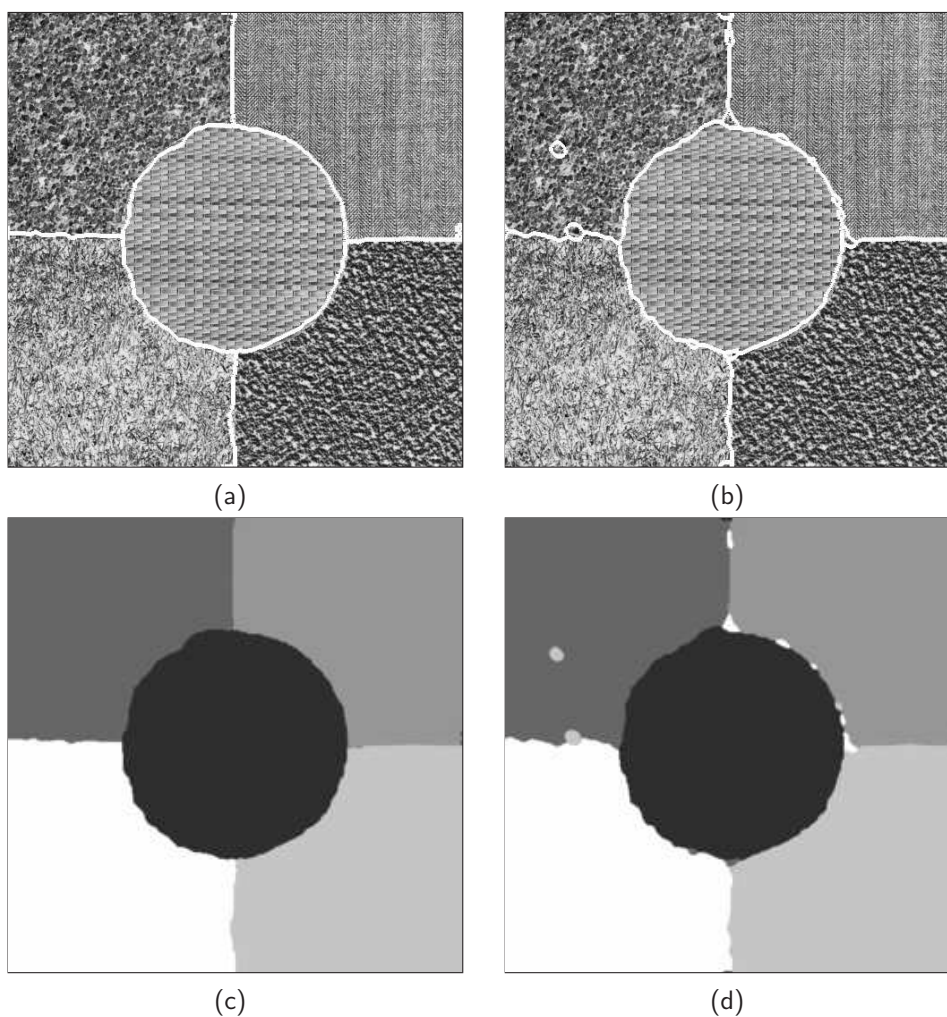


Figura 4.14: Segmentação resultante para o mosaico #1, composto de 512×512 pixels, contendo 5 classes distintas. (a) ICM com aplicação da perturbação, apresentando segmentação correta de 98.4%; (b) ICM sem aplicação da perturbação, a segmentação correta fica em 97.9%; (c)-(d) realizações resultantes do MRF.

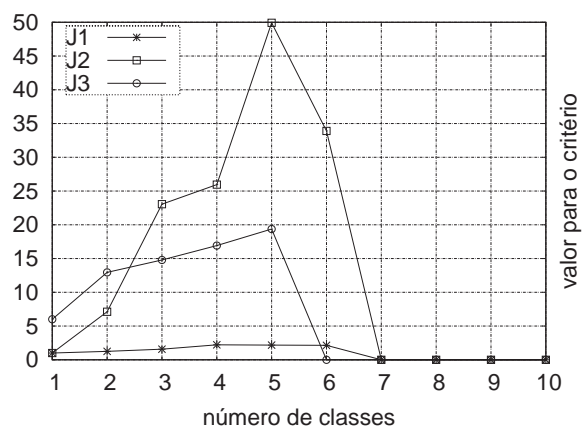


Figura 4.15: Critérios para determinação do número de classes presentes na imagem da figura 4.14.

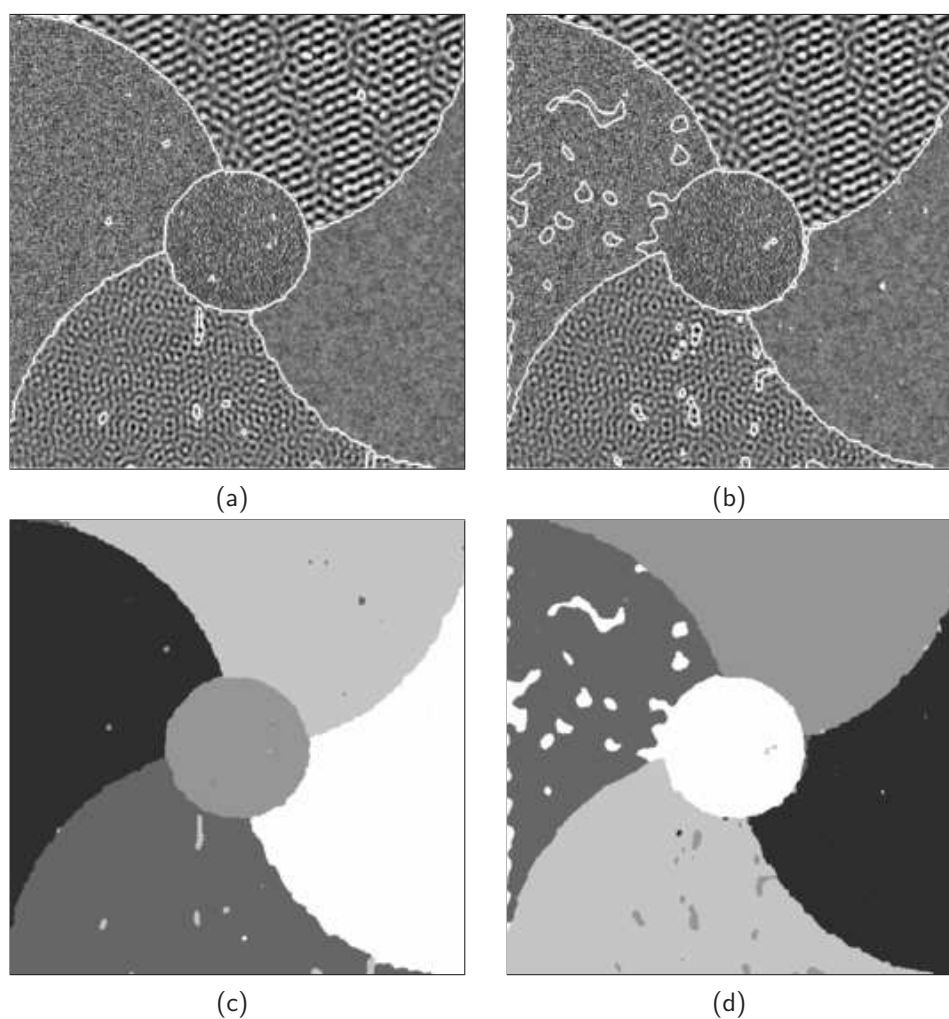


Figura 4.16: Segmentação resultante para o mosaico #2, composto de 512×512 pixels, contendo 5 classes distintas. (a) ICM com aplicação da perturbação, apresentando segmentação correta de 98.3%; (b) ICM sem aplicação da perturbação, a segmentação correta fica em 95.9%; (c)-(d) realizações resultantes do MRF.

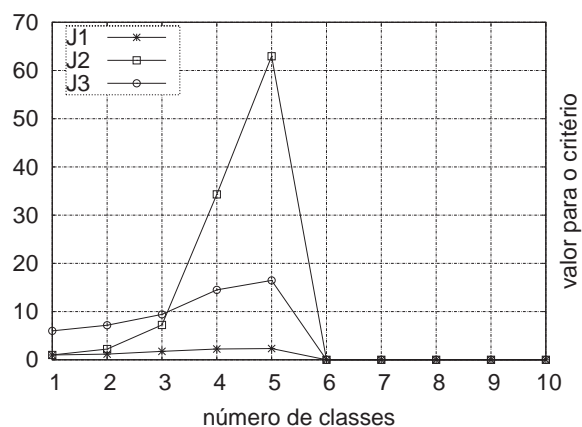


Figura 4.17: Critérios para determinação do número de classes presentes na imagem da figura 4.16.

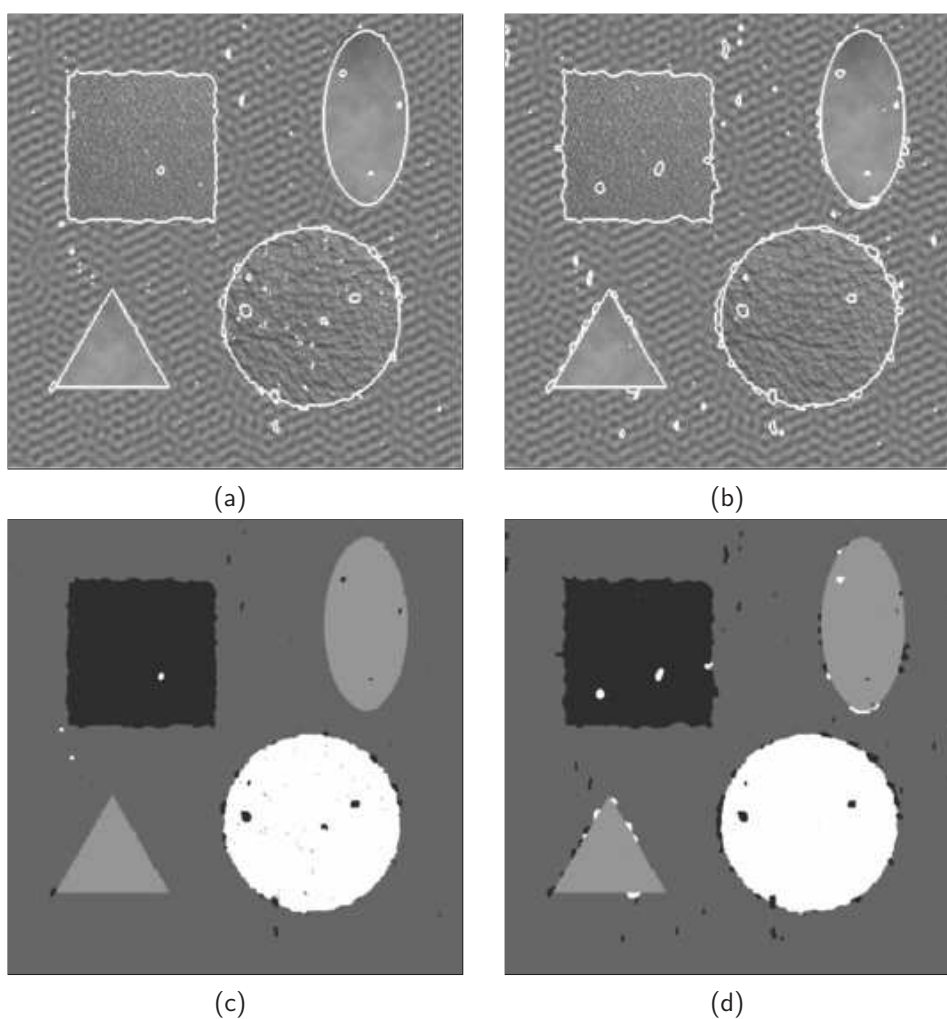


Figura 4.18: Segmentação resultante para o mosaico #3, composto de 512×512 pixels, contendo 4 classes distintas, embora dispostas em 5 regiões. (a) ICM com aplicação da perturbação, apresentando segmentação correta de 97.5%; (b) ICM sem aplicação da perturbação, a segmentação correta fica em 96.1%; (c)-(d) realizações resultantes do MRF.

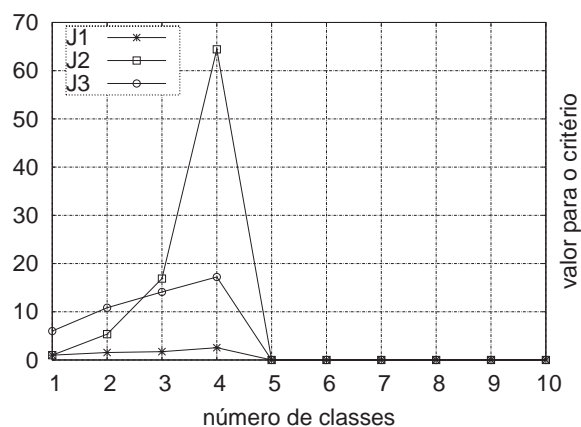


Figura 4.19: Critérios para determinação do número de classes presentes na imagem da figura 4.18.

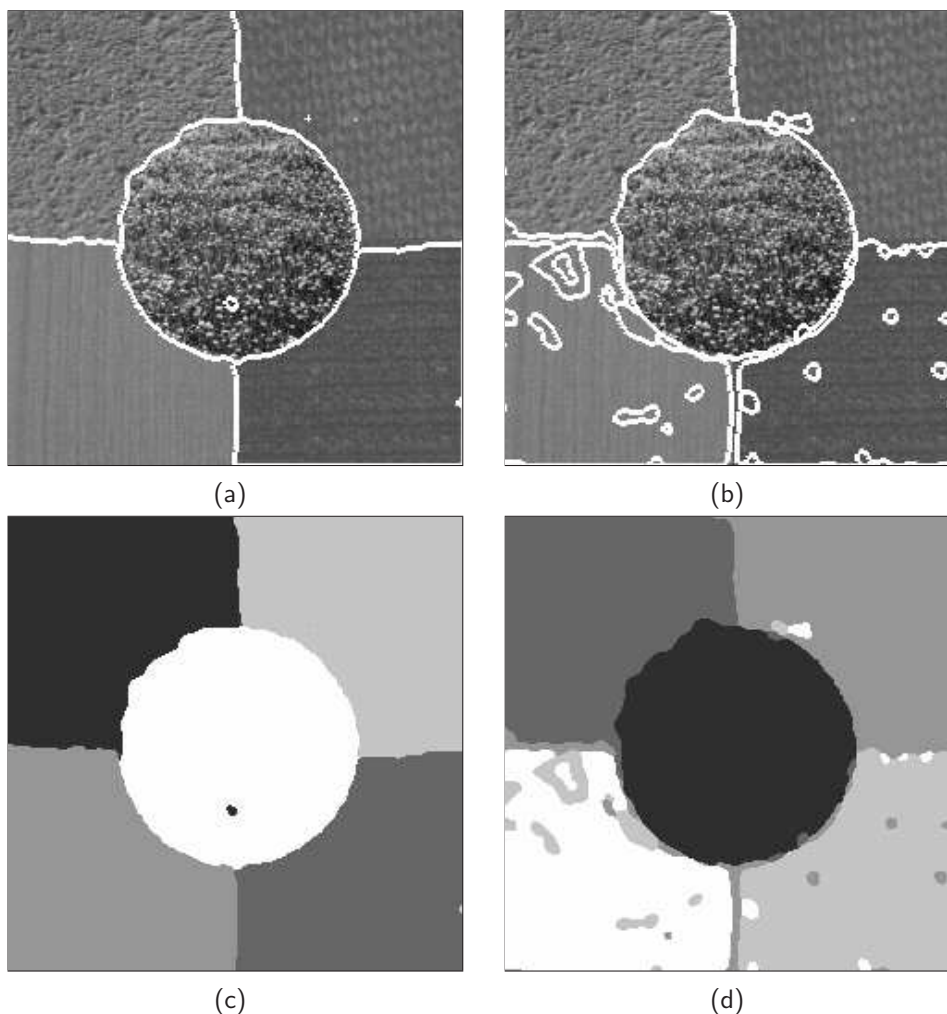


Figura 4.20: Segmentação resultante para o mosaico #4, composto de 256×256 pixels, contendo 5 classes distintas. (a) ICM com aplicação da perturbação, apresentando segmentação correta de 98.2%; (b) ICM sem aplicação da perturbação, a segmentação correta fica em 93.9%; (c)-(d) realizações resultantes do MRF.

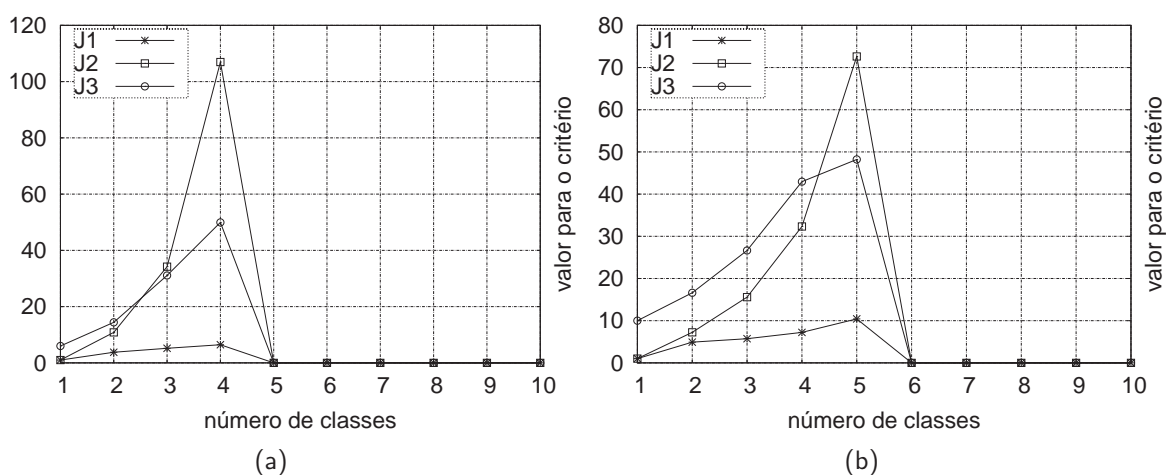


Figura 4.21: Critérios para determinação do número de classes presentes na imagem da figura 4.20. (a) utilizando o método baseado em transformada wavelet; (b) utilizando método baseado em matrizes de comprimento de corridas de cinza.

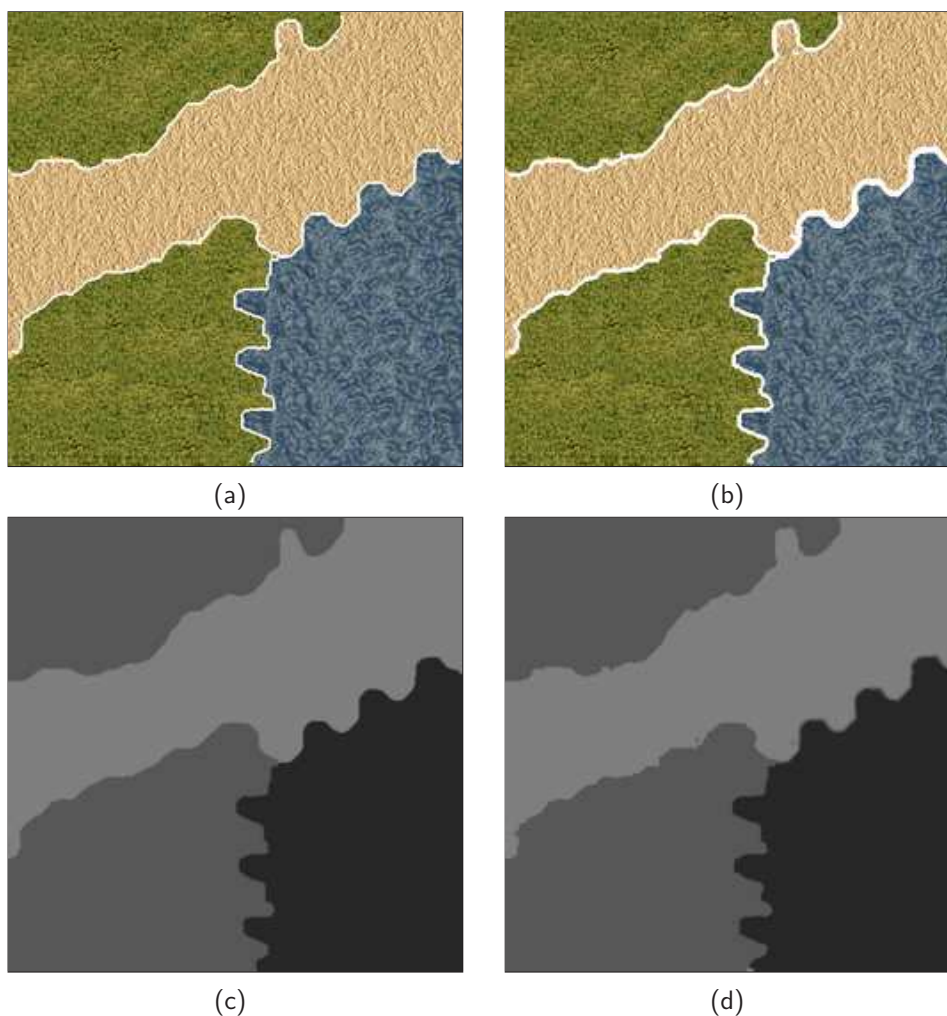


Figura 4.22: Segmentação resultante para imagem composta de 512×512 pixels. (a) resultado considerando a dependência espacial; (b) resultado utilizando apenas o classificador quadrático; (c) realização resultante do MRF, considerando a dependência espacial; (d) regiões resultantes da segmentação baseada no classificador quadrático.

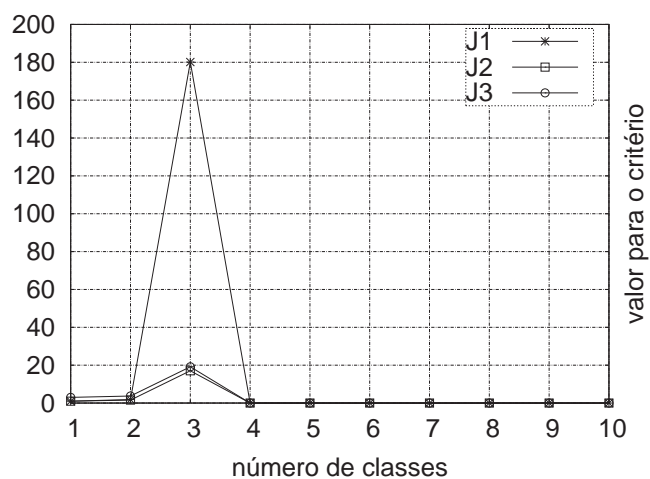


Figura 4.23: Critérios para determinação do número de classes presentes na imagem da figura 4.22.



Figura 4.24: Segmentação resultante para imagem composta de 512×256 pixels.

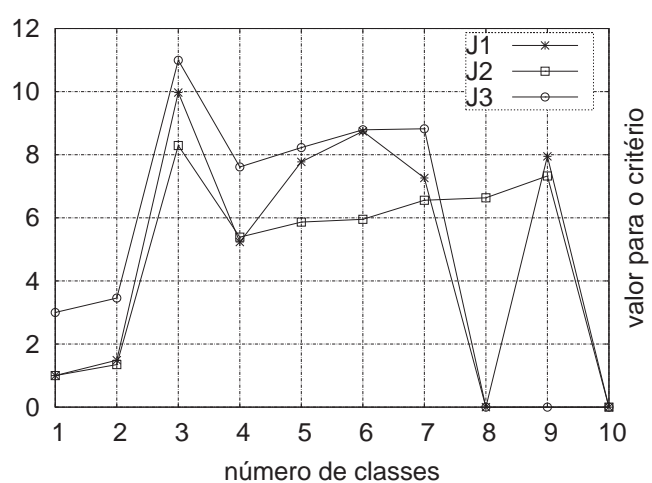


Figura 4.25: Critérios para determinação do número de classes presentes na imagem da figura 4.24.

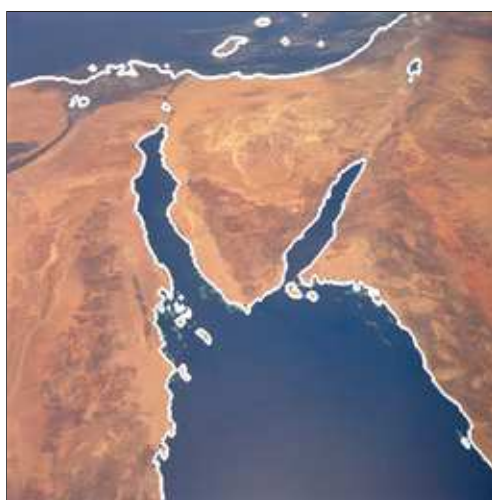


Figura 4.26: Segmentação resultante para imagem composta de 640×640 pixels.

tivo de permitir que estados de maior energia sejam alcançados de modo a reduzir a influência das condições iniciais. A perturbação é considerada nos experimentos cujos resultados são mostrados nas figuras 4.14, 4.16, 4.18 e 4.20, nos quais utiliza-se a realização inicial aleatória para o MRF (detalhes na seção 3.3).

Por meio dos resultados obtidos percebe-se que a segmentação resultante é mais acurada quando se aplica a perturbação. Adicionalmente, os resultados se mostram robustos quanto às condições iniciais, ou seja, a realização inicial aleatória para o MRF apresenta pouca influência na segmentação final.

Embora aplicada a perturbação, que atua na redução de regiões isoladas, a imagem mostrada na figura 4.18(c) indica a presença dessas regiões. Isso acontece devido à influência das características dos pixels, pois a perturbação apenas atribui novas classes às variáveis aleatórias localizadas em regiões isoladas, porém, a execução do ICM continua de maneira normal. Dessa maneira, tais regiões podem ser novamente atribuídas às antigas classes.

Por outro lado, se as regiões isoladas fossem atribuídas definitivamente a uma das classes vizinhas, o problema que ocorre neste caso específico seria resolvido. Entretanto, acarretaria na eliminação de áreas estreitas presentes na imagem, por exemplo, possivelmente eliminaria rios ou penínsulas contidos em imagens de sensoriamento remoto.

Outro experimento efetuado refere-se à determinação do número de classes presentes na imagem, baseada na aplicação dos critérios J_1 , J_2 e J_3 descritos na seção 3.2. Em geral, percebe-se que os critérios J_2 e J_3 apresentam os melhores resultados, determinando corretamente o número de classes presentes nas imagens 4.14, 4.16, 4.18 e 4.22. Alguns problemas são encontrados para as imagens 4.20 e 4.24.

Na imagem 4.20, quando se utiliza o método baseado em transformada wavelet para análise de texturas são detectadas apenas quatro classes. Nos experimentos efetuados, as duas regiões situadas mais abaixo são consideradas como pertencentes a uma mesma classe. Isto ocorre porque essas regiões são diferenciadas basicamente pelas tonalidades presentes nas texturas, as quais são desconsideradas na composição do vetor de características pelo método baseado em transformada wavelet. Dessa maneira, utiliza-se o método baseado em matrizes de comprimento de corridas de cinza, que considera os tons de cinza durante a criação das matrizes, para efetuar a segmentação da imagem, obtendo-se, então, cinco classes.

Enquanto para a imagem natural, mostrada na figura 4.24, os critérios apresentam valores máximos para três classes (gráfico da figura 4.25), pelo menos cinco classes podem ser identificadas, o que se deve ao tamanho da janela utilizada. Em imagens naturais, a detecção do número de classes é uma tarefa complexa (Lau e Levine 2002), pois é dependente de quais

elementos devem ser considerados como relevantes, o que depende diretamente da escala em questão.

A influência da escala pode ser percebida na figura 4.24. Se a imagem for considerada em âmbito global, cinco classes podem ser identificadas: mar, céu, nuvens, vegetação e areia. Entretanto, se os detalhes forem considerados, percebe-se duas classes distintas de vegetação e duas categorias de nuvens. Dessa maneira, a determinação do número de classes em imagens naturais apresenta-se altamente dependente do interesse final da segmentação, ou seja, envolve muito mais a etapa de interpretação do que processamento de imagens.

A alteração proposta para as condições iniciais do algoritmo de agrupamento *k-means* é avaliada pelos gráficos da figura 4.27. Esse experimento consiste em determinar quais áreas das regiões homogêneas são corretamente segmentadas, ou seja, determinar a percentagem das regiões homogêneas segmentadas corretamente. O gráfico em (a) é obtido com o uso de variáveis padronizadas enquanto para (b) não se aplica a padronização, onde o eixo das abscissas indica o valor de N_a do algoritmo 3.1 e o eixo das ordenadas a segmentação correta. Para cada N_a , executa-se dez vezes o *k-means*, extraíndo a média dos valores obtidos.

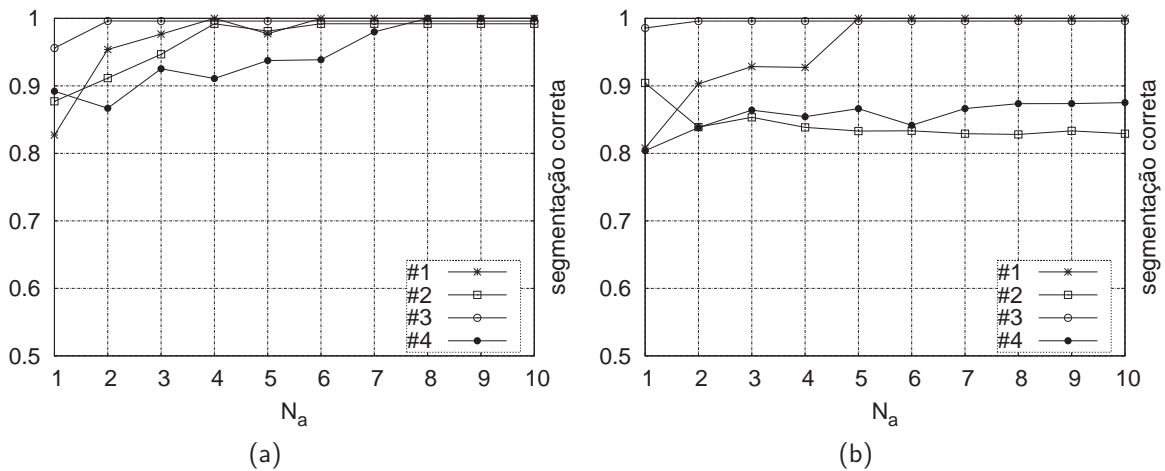


Figura 4.27: Alteração no número de execuções do *k-means*, cada curva representa um dos mosaicos apresentados anteriormente. (a) utilizando variáveis padronizadas; (b) utilizando variáveis sem padronização.

Nesse experimento, quando se utiliza variáveis padronizadas, percebe-se que o *k-means* original, $N_a = 1$, apresenta o pior desempenho, sendo instável para todos os mosaicos testados. Entretanto, conforme aumenta o valor de N_a , a instabilidade decresce e as regiões homogêneas tendem a ser corretamente segmentadas. Por outro lado, para variáveis sem padronização, os mosaicos #2 e #4 não chegam à estabilidade, inclusive apresentando resultados piores que os obtidos para valores baixos de N_a quando se utiliza a padronização.

A partir desses resultados conclui-se que a padronização é importante quando se faz ne-

cessário o cálculo de distâncias sem correção de escala, como a distância Euclidiana. Adicionalmente, a alteração efetuada na inicialização do *k-means* proporciona resultados mais estáveis conforme são aumentados os valores do parâmetro N_a , pois condições iniciais distintas podem ser avaliadas permitindo que a minimização da função de custo possa convergir para diferentes regiões de mínimos locais.

O resultado do último experimento efetuado para avaliação da metodologia proposta é apresentado no gráfico da figura 4.28, onde o eixo das abscissas indica o número de ciclos do ICM e o eixo das ordenadas representa a energia. Essa figura mostra o decaimento da energia presente no sistema, sendo apresentados 100 ciclos do ICM contendo a média da energia obtida durante 10 execuções da segunda etapa do método de segmentação para o mosaico #4, o qual apresenta maior diferença nos resultados quando se aplica a perturbação (figura 4.20).

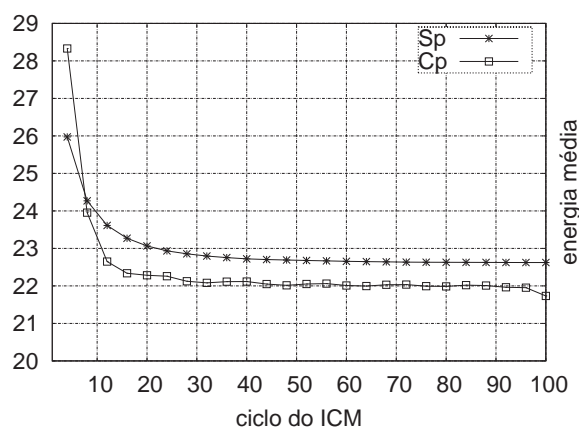


Figura 4.28: Decaimento da energia para o ICM, com perturbação (Cp) e sem perturbação (Sp). Considera-se, neste experimento, a energia total dividida pelo número de pixels contidos na imagem.

Conforme já descrito na seção 3.3, quando se aplica a perturbação, estados de maior energia podem ser obtidos. Essa característica pode ser percebida a partir do gráfico, em regiões próximas ao ciclo de número 50. Embora no início da execução a energia seja maior, logo a segmentação que utiliza a perturbação passa a apresentar menor quantidade de energia, ou seja, consegue minimizar de maneira mais intensa a função mostrada na equação 3.25.

Alguns resultados obtidos pela metodologia proposta encontram-se publicados. A proposição do histograma tridimensional e a avaliação de algumas características de texturas são realizadas por Schwartz e Pedrini (2003, 2004). A aplicação da segmentação baseada em dependência espacial para imagens monocromáticas de terrenos é descrita por Schwartz e Pedrini (2005b). Finalmente, um trabalho onde se propõe a segmentação de imagens coloridas utilizando matrizes de co-ocorrência tridimensionais em conjunto com a dependência espacial foi recentemente submetido (Schwartz e Pedrini 2005a).

CAPÍTULO 5

CONCLUSÕES

Uma etapa crítica presente no processo de análise de imagens é a segmentação, responsável por obter informações de alto nível sobre os objetos ou regiões contidos na imagem, de modo a facilitar sua interpretação. Este trabalho apresentou um novo método de segmentação de imagens coloridas e monocromáticas baseado na extração de características de texturas e na dependência espacial entre as regiões.

A utilização da dependência espacial, ao invés de apenas as características de texturas, proporciona resultados mais acurados, pois possibilita que a segmentação seja adaptável às características específicas de cada região. A divisão do método em duas etapas eliminou a necessidade de um conjunto de treinamento previamente definido. Os critérios para determinação do número de classes apresentaram resultados satisfatórios. Com uma simples alteração nas condições iniciais no *k-means*, os agrupamentos resultantes passaram a ser mais estáveis. Finalmente, os resultados apresentados validam a aplicação da perturbação no algoritmo ICM.

Com a utilização da dependência espacial, a segmentação resultante apresentou uma boa adaptação às fronteiras entre as classes tanto em imagens coloridas quanto monocromáticas. Como o método utiliza vetores de características para descrever as regiões e o MRF considera apenas a distribuição dos rótulos das classes, outras categorias de imagens podem ser utilizadas, bastando para isso, alterar o método de extração de características de texturas.

Diversos métodos de análise de texturas a partir de uma série de experimentos foram avaliados, sendo que o método baseado em transformada wavelet apresentou os melhores resultados quando se considera o compromisso entre a qualidade e o tempo de execução requerido. O seu desempenho pôde ser corroborado pelos resultados obtidos com a aplicação da metodologia proposta.

Como trabalhos futuros, o autor pretende desenvolver uma otimização para a segunda etapa da segmentação, visando à aplicação do ICM apenas em regiões consideradas heterogêneas, reduzindo dessa maneira, o custo computacional requerido. Entretanto, tal abordagem pode apresentar alguns problemas, como a eliminação de regiões estreitas durante a amostragem inicial da imagem. Uma possível solução para isso pode ser a aplicação de um desbaste nas regiões homogêneas.

Como outra futura proposição está a avaliação dos métodos de análise de texturas aqui

descritos para imagens digitais de terrenos, efetuando-se os mesmos experimentos apresentados, entretanto, sobre imagens naturais obtidas por meio de satélites. Devido ao grande interesse da comunidade de sensoriamento remoto em métodos de análise e processamento de imagens, tal avaliação pode apresentar uma boa contribuição.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abend, K., Harley, T. J. e Kanal, L. N. (1965). Classification of Binary Random Patterns, *IEEE Transactions on Information Theory* **11**(4): 538–544.
- Al-Janobi, A. (2001). Performance Evaluation of Cross-diagonal Texture Matrix Method of Texture Analysis, *Pattern Recognition* **34**(1): 171–180.
- Albregtsen, F., Nielsen, B. e Danielsen, H. E. (2000). Adaptive Gray Level Run Length Features from Class Distance Matrices, *International Conference on Pattern Recognition*, Vol. 3, pp. 738–741.
- Andrey, P. e Tarroux, P. (1998). Unsupervised Segmentation of Markov Random Field Modeled Textured Images Using Selectionist Relaxation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **20**(3): 252–262.
- Balan, A. G. R. (2003). *Técnicas de Segmentação de Imagens Aéreas para Contagem de População de Aves*, Dissertação de Mestrado, Universidade de São Paulo, São Carlos.
- Baraldi, A. e Parmiggiani, F. (1995). An Investigation of the Textural Characteristics Associated with Gray Level Cooccurrence Matrix Statistical Parameters, *IEEE Transactions on Geoscience and Remote Sensing* **33**(2): 293–304.
- Bartlett, M. S. (1968). A Further Note on Nearest Neighbour Models, *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **131**: 579–580.
- Besag, J. E. (1972). Nearest-Neighbour Systems and the Auto-Logistic Model for Binary Data, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **34**(1): 75–83.
- Besag, J. E. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **36**(2): 192–236.
- Besag, J. E. (1986). On the Statistical Analysis of Dirty Pictures, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **48**(3): 259–302.
- Blake, A. (1989). Comparison of the Efficiency of Deterministic and Stochastic Algorithms for Visual Reconstruction, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(1): 2–12.
- Brodatz, P. (1966). *Textures: A Photographic Album for Artists and Designers*, Dover Publications, Inc., New York, Estados Unidos.
- Bruno, O. M. e Costa, L. F. (2000). Effective Image Segmentation with Flexible ICM-Based Markov Random Fields in Distributed Systems of Personal Computers, *Real-Time Imaging* **6**(4): 283–295.
- Cai, J. e Liu, Z.-Q. (2002). Pattern Recognition Using Markov Random Field Models, *Pattern Recognition* **35**(3): 725–733.
- Chang, T. e Kuo, C.-C. J. (1993). Texture Analysis and Classification with Tree-Structured Wavelet Transform, *IEEE Transactions on Image Processing* **2**(4): 429–441.
- Chen, C.-C. e Huang, C.-L. (1993). Markov Random Fields for Texture Classification, *Pattern Recognition Letters* **14**(11): 907–914.

- Cheng, H. D., Jiang, X. H., Sun, Y. e Wang, J. (2001). Color Image Segmentation: Advances and Prospects, *Pattern Recognition* **34**(12): 2259–2281.
- Chu, A., Sehgal, C. M. e Greenleaf, J. F. (1990). Use of Gray Value Distribution of Run Lengths for Texture Analysis, *Pattern Recognition Letters* **11**: 415–420.
- Cohen, F. S. e Cooper, D. B. (1987). Simple Parallel Hierarchical and Relaxation Algorithms for Segmenting Noncausal Markovian Random Fields, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **9**(2): 195–219.
- Connors, R. W. e Harlow, C. A. (1980). A Theoretical Comparison of Texture Algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2**(3): 204–221.
- Cross, G. R. e Jain, A. K. (1983). Markov Random Field Texture Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **5**(1): 25–39.
- Dacheng, T., Xuelong, L., Yuan, Y., Nenghai, Y., Zhengkai, L. e Xiao-ou, T. (2002). A Set of Novel Textural Features Based on 3D Co-Occurrence Matrix for Content-Based Image Retrieval, *International Conference on Information Fusion*, Vol. 2, pp. 1403–1407.
- Daily, M. J. (1989). Color Image Segmentation Using Markov Random Fields, *IEEE Computer Society Conference in Computer Vision and Pattern Recognition*, pp. 304–312.
- de Wouwer, G. V., Scheunders, P. e D. Van Dyck, D. (1999). Statistical Texture Characterization from Discrete Wavelet Representations, *IEEE Transactions on Image Processing* **8**(4): 592–598.
- D’Elia, C., Poggi, G. e Scarpa, G. (2003). A Tree-Structured Markov Random Field Model for Bayesian Image Segmentation, *IEEE Transactions on Image Processing* **12**(10): 1259–1273.
- Dempster, A. P., Laird, N. M. e Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data Via the EM Algorithm, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **37**(1): 1–37.
- Deng, H. e Clausi, D. A. (2004). Unsupervised Image Segmentation Using a Simple MRF Model with a New Implementation Scheme, *Pattern Recognition* **37**(12): 2323–2335.
- Deng, Y. e Manjunath, B. S. (2001). Unsupervised Segmentation of Color-Texture Regions in Images and Video, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**(8): 800–810.
- Derin, H. e Elliott, H. (1987). Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **9**(1): 39–55.
- Dubes, R. C. e Jain, A. K. (1989). Random Field Models in Image Analysis, *Journal of Applied Statistics* **16**(2): 131–164.
- Duda, R. O. e Hart, P. E. (1973). *Pattern Classification and Scene Analysis*, Wiley-Interscience Publication, New York, NY, Estados Unidos.
- Farb, G. L. G. e Jain, A. K. (1996). On Retrieving Textured Images from an Image Database, *Pattern Recognition* **29**(9): 1461–1483.
- Fjortoft, R., Delignon, Y., Pieczynski, W., Sigelle, M. e Tupin, F. (2003). Unsupervised Classification of Radar Images Using Hidden Markov Chains and Hidden Markov Random Fields, *IEEE Transactions on Geoscience and Remote Sensing* **41**(3): 675–686.

- Fwu, J. K. e Djuric, P. M. (1996). Unsupervised Vector Image Segmentation by a Tree Structure-ICM Algorithm, *IEEE Transactions on Medical Imaging* **15**(6): 871–880.
- Galloway, M. M. (1975). Texture Analysis Using Gray Level Run Lengths, *Computer Graphics and Image Processing* **4**: 172–179.
- Geiger, D. e Giasi, F. (1991). Parallel and Deterministic Algorithms from MRF's: Surface Reconstruction, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(5): 401–412.
- Geman, D., Geman, S., Graffigne, C. e Dong, P. (1990). Boundary Detection by Constrained Optimization, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(7): 609–628.
- Geman, S. e Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**(6): 721–741.
- Gonzalez, R. C. e Woods, R. E. (2000). *Processamento de Imagens Digitais*, Edgard Blücher, São Paulo, SP.
- Haralick, R. M. (1979). Statistical and Structural Approaches to Texture, *Proceedings of the IEEE* **67**(5): 786–804.
- Haralick, R. M., Shanmugam, K. e Dinstein, I. (1973). Textural Features for Image Classification, *IEEE Transactions on Systems, Man and Cybernetics* **3**(6): 610–621.
- He, D. C. e Wang, L. (1990). Texture Unit, Texture Spectrum, and Texture Analysis, *IEEE Transactions on Geoscience and Remote Sensing* **28**(4): 509–512.
- He, D. C. e Wang, L. (1991). Texture Features Based on Texture Spectrum, *Pattern Recognition* **24**(5): 391–399.
- Hohn, F. E. (1958). *Elementary Matrix Algebra*, The Macmillan Company.
- Horng, M.-H., Sunb, Y.-N. e Lin, X.-Z. (2002). Texture Feature Coding Method for Classification of Liver Sonography, *Computerized Medical Imaging and Graphics* **26**(1): 33–42.
- Huang, C.-L., Cheng, T.-Y. e Chen, C.-C. (1992). Color Images Segmentation Using Scale Space Filter and Markov Random Field, *Pattern Recognition* **25**(10): 1217–1229.
- Jackson, Q. e Landgrebe, D. A. (2002). Adaptive Bayesian Contextual Classification Based on Markov Random Fields, *IEEE Transactions on Geoscience and Remote Sensing* **40**(11): 2454–2463.
- Kanal, L. N. (1980). Markov Mesh Models, *Computer Graphics and Image Processing* **12**(4): 371–375.
- Kervrann, C. e Heitz, F. (1995). A Markov Random Field Model-Based Approach to Unsupervised Texture Segmentation Using Local and Global Spatial Statistics, *IEEE Transactions on Image Processing* **4**(6): 856–862.
- Kim, J. H., Yun, I. D. e Lee, S. U. (1998). Unsupervised Segmentation of Textured Image Using Markov Random Field in Random Spatial Interaction, *International Conference on Image Processing*, Vol. 3, pp. 756–760.
- Kindermann, R. e Snell, J. L. (1980). *Markov Random Fields and their Applications*, Contemporary Mathematics, American Mathematical Society.

- Kirkpatrick, S., Gelatt, C. D. e Vecchi, M. P. (1983). Optimization by Simulated Annealing, *Science* **220**: 671–680.
- Lau, H. F. e Levine, M. D. (2002). Finding a Small Number of Regions in an Image Using Low-Level Features, *Pattern Recognition* **35**(11): 2323–2339.
- Liapis, S., Sifakis, E. e Tziritas, G. (2004). Colour and Texture Segmentation Using Wavelet Frame Analysis, Deterministic Relaxation, and Fast Marching Algorithms, *Journal of Visual Communication and Image Representation* **15**(1): 1–26.
- Loh, H.-H., Leu, J.-G. e Luo, R. C. (1988). The Analysis of Natural Textures Using Run Length Features, *IEEE Transactions on Industrial Electronics* **35**(2): 323–328.
- Lucchese, L. e Mitra, S. (2001). Color Image Segmentation: A State-of-the-Art Survey, *Proceedings of the Indian National Science Academy* **67**(2): 207–221.
- Luthon, F., Caplier, A. e Liévin, M. (1999). Spatiotemporal MRF Approach to Video Segmentation: Application to Motion Detection and Lip Segmentation, *Signal Processing* **76**(1): 61–80.
- Mallat, S. G. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(7): 674–693.
- Manjunath, B. e Chellappa, R. (1991). Unsupervised Texture Segmentation Using Markov Random Field Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(5): 478–482.
- Manjunath, B. S., Simchony, T. e Chellappa, R. (1990). Stochastic and Deterministic Networks for Texture Segmentation, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **38**(6): 1039–1049.
- Marroquin, J., Mitter, S. e Poggio, T. (1987). Probabilistic Solution of Ill-posed Problems in Computational Vision, *Journal of the American Statistical Association* **82**(397): 76–89.
- MeasTex (2005). MeasTex Texture Database. <http://www.cssip.uq.edu.au/meastex/www/obtaining.html>.
- Metropolis, N., Rosenbluth, A. W., M. N. Rosenbluth, A. H. T. e Teller, E. (1953). Equation of State Calculation by Fast Computing Machines, *Journal of Chemical Physics* **21**(6): 1087–1091.
- Milsztajn, F. (2003). *Segmentação de Tecidos Cerebrais em Imagens de Ressonância Magnética Utilizando Campos Aleatórios de Markov*, Dissertação de Mestrado, Universidade Federal do Paraná, Curitiba.
- Ohanian, P. P. e Dubes, R. C. (1992). Performance Evaluation for Four Classes of Textural Features, *Pattern Recognition* **25**(8): 819–833.
- Ojala, T., Pietikäinen, M. e Harwood, D. (1996). A Comparative Study of Texture Measures with Classification Based on Featured Distributions, *Pattern Recognition* **29**(1): 51–59.
- Pal, K. R. e Pal, S. K. (1993). A Review On Image Segmentation Techniques, *Pattern Recognition* **26**(9): 1277–1294.
- Palm, C. (2004). Color Texture Classification by Integrative Co-Occurrence Matrices, *Pattern Recognition* **37**(5): 965–976.

- Panjwani, D. K. e Healey, G. (1995). Markov Random Field Models for Unsupervised Segmentation of Textured Color Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(10): 939–954.
- Pickard, D. K. (1980). Unilateral Markov Random Fields, *Advanced Applied Probability* **12**: 655–671.
- Pietikäinen, M., Ojala, T. e Xu, Z. (2000). Rotation-Invariant Texture Classification Using Feature Distributions, *Pattern Recognition* **33**(1): 43–52.
- Poggi, G. e Ragozini, A. R. P. (1999). Image Segmentation by Tree-Structured Markov Random Field, *IEEE Signal Processing Letters* **6**(7): 155–157.
- Rajapakse, J. C., Giedd, J. N. e Rapoport, J. L. (1997). Statistical Approach to Segmentation of Single-Channel Cerebral MR Images, *IEEE Transactions on Medical Imaging* **16**(2): 176–186.
- Randen, T. e Husoy, J. H. (1999). Filtering for Texture Classification: A Comparative Study, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(4): 291–310.
- Reed, T. e Dubuf, J. M. H. (1993). A Review of Recent Texture Segmentation and Feature Extraction Techniques, *CVGIP: Image Understanding* **57**(3): 359–372.
- Rosenfeld, A. e Troy, E. B. (1970). Visual Texture Analysis, *Conf. Record for Symposium on Feature Extraction and Selection in Pattern Recognition*, pp. 115–124.
- Rosholm, A. (1997). *Statistical Methods for Segmentation and Classification of Images*, Tese de Doutorado, Technical University of Denmark.
- Ruiz, L. A., Fdez-sarría, A. e Recio, J. A. (2004). Texture Feature Extraction for Classification of Remote Sensing Data Using Wavelet Decomposition: A Comparative Study, *Proceedings International Society for Photogrammetry and Remote Sensing Congress*, Istambul, Turquia.
- Schwartz, W. R. e Pedrini, H. (2003). Método para Classificação de Imagens Baseada em Matrizes de Co-ocorrência Utilizando Características de Textura, *III Colóquio Brasileiro de Ciências Geodésicas*, Curitiba-PR, Brasil, pp. 1–11.
- Schwartz, W. R. e Pedrini, H. (2004). Texture Classification Based on Spatial Dependence Features Using Co-Occurrence Matrices and Markov Random Fields, *IEEE International Conference on Image Processing*, Vol. 1, Cingapura, pp. 239–242.
- Schwartz, W. R. e Pedrini, H. (2005a). Color Texture Image Segmentation Based on Spatial Dependence Using 3D Co-Occurrence Matrices and Markov Random Fields, *Submetido para IEEE International Conference on Image Processing*, Gênova, Itália.
- Schwartz, W. R. e Pedrini, H. (2005b). Segmentação de Imagens de Terrenos Baseada na Associação de Características de Texturas com Dependência Espacial Modelada por Campo Aleatório de Markov, *Aceito para apresentação no XII Simpósio Brasileiro de Sensoriamento Remoto*, Goiânia-GO.
- Sethian, J. (1996). A Fast Marching Level Set Method for Monotonically Advancing Fronts, *Proceedings of the National Academy of Sciences of the United States of America* **93**(4): 1591–1595.
- Sharma, M. (2001). *Performance Evaluation of Image Segmentation and Texture Extraction Methods in Scene Analysis*, Dissertação de Mestrado, University of Exeter, Reino Unido.

- Simchony, T. e Chellappa, R. (1988). Stochastic and Deterministic Algorithms for MAP Texture Segmentation, *International Conference on Acoustics, Speech, and Signal Processing*, Vol. 2, pp. 1120–1123.
- Singh, M. e Singh, S. (2002). Spatial Texture Analysis: A Comparative Study, *International Conference on Pattern Recognition*, Vol. 1, pp. 676–679.
- Smith, G. e Burns, I. (1997). Measuring Texture Classification Algorithms, *Pattern Recognition Letters* **18**(14): 1495–1501.
- Stewart, D., Blacknell, D., Blake, A., Cook, R. e Oliver, C. (2000). Optimal Approach to SAR Image Segmentation and Classification, *IEE Proceedings - Radar, Sonar and Navigation* **147**(3): 134–142.
- Tamura, H., S. M. e Yamawaki, Y. (1978). Textural Features Corresponding to Visual Preception, *IEEE Transactions on Systems, Man, and Cybernetics* **8**: 237–247.
- Tang, X. (1998). Texture Information in Run-Length Matrices, *IEEE Transactions on Image Processing* **7**(11): 1602–1609.
- Theodoridis, S. e Koutroumbas, K. (2003). *Pattern Recognition*, 2^a edn, Elsevier Academic Press, San Diego, Estados Unidos.
- Tonazzini, A. e Bedini, L. (2003). Monte Carlo Markov Chain Techniques for Unsupervised MRF-Based Image Denoising, *Pattern Recognition Letters* **24**(1): 55–64.
- Tseng, D.-C. e Lai, C.-C. (1999). A Genetic Algorithm for MRF-based Segmentation of Multi-spectral Textured Images, *Pattern Recognition Letters* **20**(14): 1499–1510.
- Tuceryan, M. e Jain, A. K. (1998). Texture Analysis, in C. H. Chen, L. F. Pau e P. Wang (eds), *The Handbook of Pattern Recognition and Computer Vision*, World Scientific Publishing Co., pp. 207–248.
- Wesolkowski, S. e Fieguth, P. (2002). Adaptive Color Image Segmentation Using Markov Random Fields, *International Conference on Image Processing*, Vol. 3, pp. 769–772.
- Weszka, J., Dyer, C. e Rosenfeld, A. (1976). A Comparative Study of Texture Measures for Terrain Classification, *IEEE Transactions on System, Man, and Cybernetics* **6**(4): 269–285.
- Winkler, G. (2003). *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods, A Mathematical Introduction*, Application of Mathematics, Stochastic Modeling and Applied Probability, Springer.
- Yong, H. (2002). *Texture Analysis by Statistical Methods*, Tese de Doutorado, Nanyang Technological University.
- Zhang, J. e Tan, T. (2002). Brief Review of Invariant Texture Analysis Methods, *Pattern Recognition* **35**(3): 735–747.